

Fusion peptides with binding activity for streptavidin

Patent Number: ☐ US5506121
Publication date: 1996-04-09
Inventor(s): SKERRA ARNE (DE); SCHMIDT THOMAS (DE)
Applicant(s):: INST BIOANALYTIK GEMEINNUETZIG (DE)
Requested Patent: ☐ DE4237113
Application Number: US19930148675 19931103
Priority Number(s): DE19924237113 19921103
IPC Classification: C07K19/00 ; C12N15/62
EC Classification: C12N15/62
Equivalents: ☐ FR2697525, ☐ GB2272698, ☐ JP7076596

Abstract

A peptide according to the invention comprises the amino acid sequence Trp-X-His-Pro-Gln-Phe-Y-Z, in which X represents any desired amino acid and Y and Z either both denote Gly, or Y denotes Glu and Z denotes Arg or Lys. A fusion protein according to the invention consists of the amino acid sequence of a complete protein, of a protein mutant such as a deletion mutant or substitution mutant, or of part of a protein linked to the sequence of a peptide of the invention. For the production of a recombinant protein by expression of a DNA sequence coding therefor in suitable host cells according to well-known methods, a DNA sequence is used which codes for a fusion protein and, if desired, the presence of the expression product is detected by means of a conjugate of streptavidin and a label or the desired protein is separated as a fusion protein by means of streptavidin affinity chromatography.

Data supplied from the esp@cenet database - I2



19 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

12 Offenlegungsschrift
10 DE 42 37 113 A 1

21 Aktenzeichen: P 42 37 113.9
22 Anmeldetag: 3. 11. 92
43 Offenlegungstag: 5. 5. 94

51 Int. Cl.⁵:
C 07 K 7/06
C 07 K 15/04
C 07 K 3/20
C 12 N 15/63
C 12 Q 1/42
G 01 N 33/68
// C12N 15/62, 15/70
(C12N 1/21, C12R
1:19)

DE 42 37 113 A 1

71 Anmelder:

Klaus Kühn Konstruktion GmbH & Co. KG, 12349
Berlin, DE

74 Vertreter:

Weickmann, H., Dipl.-Ing.; Fincke, K., Dipl.-Phys.
Dr.; Weickmann, F., Dipl.-Ing.; Huber, B.,
Dipl.-Chem.; Liska, H., Dipl.-Ing. Dr.-Ing.; Prechtel,
J., Dipl.-Phys. Dr.rer.nat.; Böhm, B., Dipl.-Chem.Univ.
Dr.rer.nat., Pat.-Anwälte, 81679 München

72 Erfinder:

Skerra, Arne, Dipl.-Ing. Dr.rer.nat., 6200 Wiesbaden,
DE; Schmidt, Thomas, Dipl.-Biol., 6382
Friedrichsdorf, DE

54 Fusionspeptide mit Bindungsaktivität für Streptavidin

67 Ein erfindungsgemäßes Peptid umfaßt die Aminosäuresequenz Trp - X - His - Pro - Gln - Phe - Y - Z, wobei X eine beliebige Aminosäure darstellt und Y und Z entweder beide Gly oder Y Glu und Z Arg oder Lys bedeuten. Ein erfindungsgemäßes Fusionsprotein besteht aus der Aminosäuresequenz eines vollständigen Proteins, einer Protein-Mutante, wie einer Deletions- oder Substitutions-Mutante, oder einem Proteinteil verbunden mit der Sequenz des Peptids nach Anspruch 1. Zur Herstellung eines rekombinanten Proteins durch Expression einer dafür kodierenden DNA-Sequenz in geeigneten Wirtszellen nach an sich bekannten Methoden verwendet man eine DNA-Sequenz, welche für ein Fusionsprotein gemäß Anspruch 2 kodiert und weist gegebenenfalls die Anwesenheit des Expressionsproduktes über ein Konjugat von Streptavidin und einer Markierung nach bzw. bewirkt die Abtrennung des gewünschten Proteins als Fusionsprotein über eine Streptavidinaffinitätschromatographie.

DE 42 37 113 A 1

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

BUNDESDRUCKEREI 03. 94 408 018/416

27/93

Beschreibung

Die vorliegende Erfindung betrifft Peptide, welche einem Protein, das mit einem solchen Peptid fusioniert ist, eine Bindungsaffinität für Streptavidin verleihen, Fusionsproteine, die ein solches Peptid enthalten sowie Verfahren zur Herstellung eines derartigen rekombinanten Proteins durch Expression einer dafür kodierenden DNA-Sequenz in geeigneten Wirtszellen nach an sich bekannten Methoden.

Während Systeme zur effizienten Expression von Fremdproteinen in einer Vielzahl von verschiedenen Wirtsorganismen weitgehend etabliert sind, stellen der Nachweis und die Reinigung des rekombinanten Genprodukts bei der gentechnologischen Proteinherstellung nach wie vor ein Problem dar. Dies gilt vor allem dann, wenn das natürliche Protein von Interesse nicht zuvor isoliert worden ist, so daß es weder hinsichtlich seiner biochemischen Eigenschaften charakterisiert werden konnte, noch spezifische Antiseren oder monoklonale Antikörper gegen das Protein erhältlich sind. Diese Situation ist in letzter Zeit aufgrund der Entwicklung der Polymerasekettenreaktion gehäuft aufgetreten (Bloch W. (1991) *Biochemistry* 30, 2736—2747), weil es dadurch oft leichter ist, Informationen über ein das Protein kodierendes Gen zu erhalten, als das gereinigte natürliche Protein selbst.

Um den Nachweis eines rekombinanten Genprodukts in solchen Fällen zu ermöglichen, wo spezifische Immunreagenzien für das Protein nicht verfügbar sind, wurden kurze Peptid-Anhängsel (Peptid-tags) verwendet, die mit dem rekombinanten Protein auf Genebene fusioniert wurden. Solche Fusionspeptidsequenzen werden von spezifischen Antikörpern erkannt, wenn sie am amino- oder carboxyterminalen Ende einer Proteinsequenz angefügt werden. Beispiele für derartige Peptid-Anhängsel sind das Myc-tag (Munro & Pelham (1986) *Cell* 46, 291—300; Ward et al. (1989) *Nature* 341, 544—546), das Flag-Peptid (Hopp et al. (1988) *Bio/Technology* 6, 1204—1210), das KT3-Epitop-Peptid (Martin et al. (1990) *Cell* 63, 843—849; Martin et al. (1992) *Science* 255, 192—194), ein α -Tubulinepitop-Peptid (Skinner et al. (1991) *J. Biol. Chem.* 266, 14163—14166) und das T7 Gen 10-Protein-Peptid-Tag (Lutz-Freyermuth et al. (1990) *Proc. Natl. Acad. Sci. USA* 87, 6393—6397), welche erfolgreich für den Nachweis und in einigen Fällen auch für die Reinigung des rekombinanten Genprodukts verwendet wurden. Zusätzlich wurde in den meisten der genannten Beispiele festgestellt, daß diese kurzen Peptid-Anhängsel, welche normalerweise 3 bis 12 Aminosäuren lang sind, nicht mit der biologischen Funktion des Proteins interferieren und deshalb nicht unbedingt nach der Expression abgespalten werden müssen.

Obwohl diese Fusionspeptide vorzüglich geeignet sind für die Detektion eines Proteins und deshalb wichtige Hilfsmittel für die Optimierung von Expressionsausbeuten und von Reinigungsprotokollen sind, sind die Anwendungsmöglichkeiten der ihnen innewohnenden Affinitätseigenschaften in einem Reinigungsschema begrenzt. Der Grund hierfür ist, daß die vorteilhaften Eigenschaften der bisher beschriebenen Peptide auf ihrer starken Bindung an einen Antikörper basieren. Folglich müssen, wenn das Protein über das Peptid-Anhängsel an eine Affinitätssäule gebunden ist, welche einen immobilisierten Antikörper trägt, ziemlich extreme und damit wenig schonende Bedingungen (d. h. unphysiologischer pH-Wert oder chaotrope Reagenzien) angewandt werden, um das Protein wieder von der Säule zu eluieren. Wenn das rekombinante Protein in einer funktionellen Form produziert wurde, ist es jedoch wünschenswert, potentiell denaturierende Bedingungen während der Reinigung zu vermeiden. Allerdings wurde nur für drei der oben beschriebenen Beispiele (Hopp et al. (1988); Martin et al. (1990); Skinner et al. (1991) *supra*) gezeigt, daß es möglich ist, die Proteinpeptidfusion unter Verwendung milderer Bedingungen, d. h. z. B. kompetitiv mit dem synthetischen Peptid zu eluieren. Auch in diesen Fällen verbleiben jedoch Nachteile, da die gegen die Peptid-Anhängsel gerichteten monoklonalen Antikörper entweder teuer oder schwer in ausreichenden Mengen zu erhalten sind.

Es war daher Aufgabe der vorliegenden Erfindung, kurze Peptidsequenzen zu entwickeln, die

- i) mit einem rekombinanten Protein verbunden werden können, ohne mit dessen Funktion zu interferieren,
- ii) den Nachweis mit einem leicht verfügbaren Reagenz ermöglichen und
- iii) leicht kontrollierbare Bindungseigenschaften zeigen.

Gelöst wird die Aufgabe erfindungsgemäß durch ein Peptid, umfassend die Aminosäuresequenz Trp-X-His-Pro-Gln-Phe-Y-Z, wobei X eine beliebige Aminosäure darstellt und Y und Z entweder beide Gly oder Y Glu und Z Arg oder Lys bedeuten. Im Rahmen der Erfindung wurde nämlich festgestellt, daß die angegebene Peptidsequenz eine hohe Bindungsaffinität für Streptavidin, bzw. "Kern"-Streptavidin (ein proteolytisches Spaltprodukt von Streptavidin) (Bayer, E.A., Ben-Hur, H., Hiller, Y. und Wilchek, M. (1989) *Biochem. J.* 259, 369—376) aufweist.

Wenn die angegebene Sequenz daher in einem Fusionsprotein vorhanden ist, so weist auch dieses Fusionsprotein eine hohe Affinität für Streptavidin auf. Ein entsprechendes Fusionsprotein ist daher ebenfalls Gegenstand der vorliegenden Erfindung. Die Peptidsequenz kann hierbei am carboxyterminalen Ende des Fusionsproteins vorhanden sein, kann theoretisch jedoch auch am aminoterminalen Ende oder innerhalb der Aminosäuresequenz des Proteins liegen, solange hiermit keine negativen Eigenschaften verbunden sind, wie z. B. Hinderung oder Zerstörung der biologischen Aktivität etc.

Das im Fusionsprotein vorhandene Protein kann sowohl ein vollständiges Protein, sowie eine Mutante eines Proteins, wie z. B. eine Deletionsmutante oder Substitutionsmutante sein, und schließlich ist es auch möglich, hier lediglich einen interessierenden Teil eines Proteins mit dem erfindungsgemäßen verbunden zu erhalten.

Die erfindungsgemäßen Fusionsproteine können leicht nachgewiesen werden durch Bindung an ein Konjugat aus Streptavidin und einer Markierung, wobei als Markierung sämtliche dem Fachmann bekannte Markierungen verwendet werden können. Auch der sonstige Ablauf eines solchen Proteinnachweises kann unter dem Fachmann geläufigen Bedingungen durchgeführt werden.

Ein weiterer Gegenstand der Erfindung ist ein Expressionsvektor, insbesondere ein bakterieller Expressionsvektor, der exprimierbar unter der Kontrolle eines geeigneten Promotors und Operators plazierte eine DNA-Sequenz enthält, welche ein erfindungsgemäßes Peptid kodiert, sowie mehrere Restriktionsschnittstellen in

5'-Richtung anschließend an diese DNA-Sequenz aufweist, welche die Einbringung einer weiteren DNA-Sequenz erlaubt, die für das zu exprimierende Protein oder ein Proteinteil kodiert.

Mit Hilfe des erfindungsgemäßen Expressionsvektors wird es ermöglicht, die DNA-Sequenz für ein interessierendes Protein in einfacher Weise vor eine DNA-Sequenz für das erfindungsgemäße Peptid zu plazieren und nach Expression z. B. in *Escherichia coli* also ein erfindungsgemäßes Fusionsprotein zu erhalten. Wird in die Restriktionsschnittstelle in 5'-Richtung von der Peptidsequenz die DNA-Sequenz für das Protein inseriert, was wie auch die übrigen Verfahrensmaßnahmen bei der Herstellung des erfindungsgemäßen Expressionsvektors für den Fachmann gängige Maßnahmen sind, so wird ein Fusionsprotein erhalten, welches das die Streptavidin-Affinität vermittelnde erfindungsgemäße Peptid am Carboxy-Terminus aufweist.

Die Restriktionsschnittstelle muß im erfindungsgemäßen Expressionsvektor nicht unbedingt unmittelbar neben der ersten oder letzten Base der für das Peptid kodierenden DNA-Sequenz liegen. Vorzugsweise sollte sie jedoch so liegen, daß bei Transkription das Leseraster nicht beeinträchtigt wird und eine Verbindung von nur wenigen, vorzugsweise höchstens zehn zusätzlichen Aminosäuren zwischen dem Peptid und der Aminosäuresequenz des Proteins entsteht.

Ein weiterer Gegenstand der vorliegenden Erfindung ist ein Verfahren zur Herstellung eines rekombinanten Proteins durch Expression einer dafür kodierenden DNA-Sequenz in geeigneten Wirtszellen nach an sich bekannten Methoden, bei denen man eine DNA-Sequenz exprimiert, welche für ein erfindungsgemäßes Fusionsprotein kodiert. Die Vorteile dieses Verfahrens liegen darin, daß die Anwesenheit des Expressionsprodukts über ein Konjugat von Streptavidin und einer Markierung leicht nachgewiesen werden kann oder/und die Abtrennung zur Aufreinigung des gewünschten Proteins als Fusionsprotein über eine Streptavidin-Affinitäts-Chromatographie bewirkt werden kann.

Wie bereits oben ausgeführt, kann zum Nachweis der Anwesenheit des Expressionsproduktes, also des Fusionsproteins, Streptavidin als Konjugat mit einer beliebigen Markierung verwendet werden, wobei im Rahmen der Erfindung eine Enzymmarkierung bevorzugt ist. Es können im übrigen die an sich bekannten Methoden zum Nachweis von Proteinen dabei zur Anwendung kommen, z. B. ELISA, RIA, Western-Transfer etc.

Ein weiterer Vorteil des erfindungsgemäßen Verfahrens liegt darin, daß die Aufreinigung des exprimierten Fusionsproteins leicht durch Affinitätschromatographie über eine Säule mit immobilisiertem Streptavidin durchgeführt werden kann. Die Elution kann dann vorteilhaft unter sehr milden Bedingungen durchgeführt werden, z. B. durch Zugabe von Biotin oder biotinähnlichen Verbindungen, oder auch mit Hilfe der durch Peptidsynthese gewonnenen Streptavidin-Affinitätspeptide.

Im Rahmen der vorliegenden Erfindung ist die Elution mit Biotin bevorzugt. Ebenso bevorzugt wird zur Affinitätschromatographie eine Säule verwendet, welche mit einer Streptavidin-Agarose-Matrix oder auch mit an EupergitTM gekoppeltem Streptavidin gepackt ist.

Aus obiger Darlegung der vorliegenden Erfindung ist zu ersehen, welche überragenden Vorteile dem erfindungsgemäßen Peptid und Fusionsproteinen, welche dieses Peptid enthalten, insofern innewohnen, als ein schneller und sicherer Nachweis des Expressionsproduktes Fusionsprotein ermöglicht wird, wobei zu erwähnen ist, daß Streptavidin ein billiges und leicht verfügbares Reagenz darstellt, das auch in Verbindung mit Markierungen wie Fluoreszenzmarkierung oder Enzymmarkierung erhältlich ist. Desweiteren zeigt das exprimierte Fusionsprotein leicht kontrollierbare Bindungseigenschaften an Streptavidin aufgrund der Anwesenheit des erfindungsgemäßen Peptids, so daß eine leichte Aufreinigung des Expressionsproduktes ermöglicht wird, die auch im großtechnischen Maßstab durchführbar ist.

Mit Hilfe der erfindungsgemäßen Expressionsvektoren wird die Expression eines erfindungsgemäßen Fusionsproteins erleichtert, wobei ein solcher Expressionsvektor universell für alle zu exprimierenden Protein anwendbar ist. Das erfindungsgemäße Peptid stört im Fusionsprotein die biologische Aktivität des übrigen Proteinanteils nicht und muß daher nicht unbedingt vor Weiterverwendung abgespalten werden. Sollte jedoch aus irgendwelchen Gründen eine Abspaltung erwünscht sein, so kann der erfindungsgemäße Expressionsvektor auch derart aufgebaut sein, daß er zwischen der Restriktionsschnittstelle zur Einbringung der DNA-Sequenz für das Protein und der für das Peptid kodierenden Sequenz noch eine DNA-Sequenz aufweist, welche für eine spezifische Protease-Schnittstelle kodiert. Somit wäre nach Expression und gegebenenfalls Aufreinigung oder Nachweis des Expressionsproduktes eine leichte Abspaltung der Peptidsequenz möglich.

Die folgenden Beispiele erläutern in Verbindung mit den Figuren die Erfindung weiter:

Hierbei zeigen die

Fig. 1 die DNA-Sequenz von pASK46 sowie die davon kodierten Aminosäuresequenzen in allen drei Leserastern und Angabe der singulären Restriktions-Schnittstellen;

Fig. 2 eine Restriktionskarte von pASK46;

Fig. 3 den Nachweis erfindungsgemäßer Protein-Peptidfusionen in einem Filter-Sandwich-Test mit jeweils 4 Kolonien von *E. coli*, transformiert mit verschiedenen Plasmiden;

Fig. 4 den Nachweis einer erfindungsgemäßen Protein-Peptidfusion im Western Transfer;

Fig. 5 die in einem ELISA beobachteten Bindungssignale für ein Streptavidinkonjugat mit Verdünnungsreihen der Periplasmafraktionen von induzierten *E. coli* TG1-Zellen, transformiert mit Plasmiden, die für verschiedene erfindungsgemäße Fusionen aus einem Antikörperfragment und einem Streptavidin-Affinitätspeptid kodieren;

Fig. 6A die Elutionsprofile der Streptavidin-Affinitätschromatographie mit Periplasmafraktionen, wie in Fig. 5;

Fig. 6B eine analoge Streptavidin-Affinitätschromatographie wie Fig. 6A, wobei hier das Antigen Lysozym der Periplasmafraktion zugesetzt worden war;

Fig. 7 eine Restriktionskarte von pASK60-Strep;

Fig. 8 die DNA-Sequenz von pASK60-Strep mit den davon kodierten Aminosäuresequenzen in allen drei Leserastern sowie den singulären Restriktions-Schnittstellen; und
Fig. 9 die kommentierte Sequenz des Polylinkers von pASK60-Strep.

5

Allgemeine Techniken

A) Gentechnische Methoden, Reagenzien

DNA-Manipulationen wurden nach gängigen gentechnischen Methoden (Sambrook J., Fritsch, E.F. und Maniatis T. (1989), Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) durchgeführt. Für Klonierung und Expression wurden die Stämme Escherichia coli K12 TG1 (supE, hsdΔ5, thi, Δ(lac-proAB) [F', traD36, proAB, lacI^PZΔM15]) (Gibson, T.J. (1984) Ph. D. Thesis, Cambridge University, England) und Escherichia coli K12 JM83 (ara, Δ(lac-pro AB), rpsL (= strA), Φ80, LacZΔM15) (Yanisch-Perron, C., Vieira, J. und Messing, J. (1985), Gene 33, 103–119) verwendet. Restriktionsenzyme wurden bei Boehringer Mannheim, New England Biolabs und Gibco BRL, Taq DNA Polymerase bei Promega erworben. Restriktionsverdau und Polymerase-Kettenreaktion (PCR) wurden unter den vom Hersteller empfohlenen Bedingungen durchgeführt. Für die ortsgerichtete Mutagenese wurde eine modifizierte Vorschrift nach Kunkel verwendet (Geisselsoder, J., Witney, F. und Yuckenberg, P. (1987), Biotechniques 5, 786–791). Oligodesoxynukleotid-Synthese wurde unter Verwendung eines Applied Biosystems DNA-Syntheseroboters durchgeführt. Das kovalente Streptavidin-alkalische Phosphatase-Konjugat wurde bei Amersham erworben und die Streptavidin-Agarose stammte von Biomol oder Sigma. Alle diese Reagenzien enthielten Kern-Streptavidin, eine proteolytisch verkürzte Form des Proteins (Bayer, E.A., Ben-Hur, H., Hiller, Y. und Wilchek, M. (1989) Biochem. J. 259, 369–376).

25

B) Plasmidkonstruktionen

Das in dieser Studie verwendete Plasmid pASK46 wurde aus früheren Gen-Konstrukten für die Expression des D1.3Fv-Fragmentes in E.coli (Ward E.S., Güssow, D., Griffith, A.D., Jones, P.T. und Winter, G. (1989) Nature 341, 544–546) und aus pASK40 (Skerra, A., Pfitzinger, I. und Plückthun, A. (1991) Bio/Technology, 9, 273–278) zusammengesetzt.

Die vollständige DNA-Sequenz von pASK46 ist in Fig. 1 dargestellt. Eine Restriktionskarte desselben Plasmids zeigt Fig. 2.

Es wurden sechs Derivate von pASK46 hergestellt, die für vier verschiedene Peptide am C-Terminus von VH bzw. zwei Peptide am C-Terminus von VL des D1.3 Fv-Fragments (Boulot, G., Eiselé, J.-L., Bentley, G.A., Bhat, T.N., Ward, E.S., Winter, G. und Poljak, R.J. (1990), J. Mol. Biol., 213, 617–619) kodieren.

Konstrukte mit Peptidanhängseln an VH

Die die Basenpaare 686–706 umfassende DNA-Sequenz auf pASK46 wurde durch die unten angegebenen, jeweils 30 Basenpaare langen Sequenzen, die für Nonapeptide kodieren, ersetzt.

	686	706
45 pASK46:	5'-TCCTCA-TAATAAGAGCTATGGGAGCTT-GCATGCAAATTCTA	
	SerSer-EndEnd	
pASK46-p111H:	5'-TCCTCA-GCGTGGAGGCATCCACAGTTCGGGGGCTAA-GCATGCAAATTCTA	
	SerSer-AlaTrpArgHisProGlnPheGlyGlyEnd	
50 pASK46-p141H:	5'-TCCTCA-CAGTGGCTGCATCCACAGTTCGGTGGCTAA-GCATGCAAATTCTA	
	SerSer-GlnTrpLeuHisProGlnPheGlyGlyEnd	
pASK46-p11XH:	5'-TCCTCA-GCGTGGAGGCATCCACAGTTCGAGCGCTAA-GCATGCAAATTCTA	
55	SerSer-AlaTrpArgHisProGlnPheGluArgEnd	
pASK46-p14XH:	5'-TCCTCA-CAGTGGCTGCATCCACAGTTCGAGCGCTAA-GCATGCAAATTCTA	
	SerSer-GlnTrpLeuHisProGlnPheGluArgEnd	

60

Konstrukte mit Peptidanhängseln an VL

Die die Basenpaare 1127–1169 umfassende DNA-Sequenz auf pASK46 wurde durch die unten angegebenen, 36 bzw. 44 Basenpaare langen Sequenzen ersetzt.

65

1127

1169

pASK46: 5'-ATCAAA-CGGGAACAAAACTCATCTCAGAAGAGGATCTGAATTAATAA-TGATC
IleLys-ArgGluGlnLysLeuIleSerGluGluAspLeuAsnEndEnd

pASK46-p111L: 5'-ATCAAA-TCAGCGTGGCGTCATCCACAGTTTCGGTGGCTAAGCT-TGATC
IleLys-SerAlaTrpArgHisProGlnPheGlyGlyEnd

pASK46-p11XL: 5'-ATCAAA-TCAGCGTGGCGTCATCCACAGTTTCGAGCGCTAAGCATGCAAGCT-TGATC
IleLys-SerAlaTrpArgHisProGlnPheGluArgEnd

C) Zellwachstum, Induktion und Zellaufschluß

Zur Produktion der rekombinanten Proteine wurden die mit dem entsprechenden Expressionsplasmid transformierten E. coli-Zellen bei 22°C in LB-Medium, das 100 µg/ml Ampicillin enthielt, inkubiert, bis eine OD₅₅₀ von 0,5 erreicht war. Dann wurde IPTG mit einer Endkonzentration von 1 mM zugesetzt, die Temperatur wurde ggf. auf 20°C gesenkt und die Induktion der Proteinexpression wurde 3 Std. lang fortgesetzt. Die Zellen wurden dann abzentrifugiert und in einem geeigneten Puffer resuspendiert. Für die Herstellung der periplasmatischen Zellfraktion wurden die Zellen von 1 l Kultur in 10 ml 50 mM Tris pH 8,0, 500 mM Saccharose, 1 mM EDTA resuspendiert und 30 Min. lang auf Eis inkubiert. Die Sphäroplasten wurden durch Zentrifugation sedimentiert, und der Überstand wurde vor der weiteren Verwendung sterilfiltriert. Wenn der Ansatz für die Reinigung mit einer Streptavidin-Agarose Säule vorgesehen war, wurde Avidin bis zu einer Endkonzentration von 40 µg/ml zugegeben, um freie Biotin-Gruppen zu komplexieren. Die Proteinlösung wurde durch Ultrafiltration auf ein Endvolumen von ca. 1 ml konzentriert und über Nacht gegen 1 l 50 mM Tris, pH 8,0, bei 4°C dialysiert. Geringe Mengen an Präzipitat wurden vor der Säulenchromatographie durch Zentrifugation (Microfuge, 14000 rpm, 10 Min., 4°C) oder Filtration mit einer Spin-X-Filtriereinheit (Costar) entfernt. Zur Herstellung des löslichen Teils des Gesamt-Zellproteins wurden die Zellen von 1 l Kultur in 10 ml 50 mM Tris, pH 8,0, resuspendiert und mit Hilfe eines Hochdruck-Homogenisators (French pressure cell) bei 18000 psi aufgeschlossen. Das Homogenisat wurde zentrifugiert (45000 g, 30 Min., 4°C) und dem Überstand zur Komplexierung freier Biotingruppen Avidin bis zu einer Konzentration von 40 µg/ml zugesetzt. Nach einer 30-minütigen Inkubation bei 4°C wurde die Proteinlösung sterilfiltriert und auf die Streptavidin-Agarose-Säule aufgetragen. Für die SDS-PAGE des Gesamt-Zellproteins wurden die Zellen aus 4 ml E. coli-Kultur in 400 µl 50 mM Tris, pH 8,0, resuspendiert und mit 100 µl 5× SDS-PAGE Auftrags-Puffer versetzt (siehe unten). Chromosomale DNA wurde durch Ultraschallbehandlung vor der Gel-Elektrophorese fragmentiert.

Beispiel 1

Filter-Sandwich Test

Der Filter-Sandwich Test wurde in Anlehnung an eine Strategie, die vorher von Skerra et al. (Skerra, A., Dreher, M.L. und Winter, G. (1991) Anal. Biochem. 196, 151–155) beschrieben wurde, durchgeführt. Die transformierten E. coli-Zellen wurden auf einer Nitrocellulose-Filtermembran (82 mm Durchmesser, Schleicher & Schuell) ausplattiert oder punktförmig aufgebracht, die auf einer Agar-Platte lag (mit LB-Medium, welches 100 µg/ml Ampicillin und 10 mg/ml Glucose enthielt). Die Platte wurde ungefähr 8 Std. lang bei 37°C inkubiert, bis kleine Kolonien sichtbar wurden. Parallel dazu wurde eine zweite Nitrocellulose-Filtermembran 6 Std. lang mit einer Lösung von 5 mg/ml Hühnereiweiß-Lysozym (Sigma), dem Antigen des D1.3 Antikörpers, in PBS-Puffer (4 mM KH₂PO₄, 16 mM Na₂HPO₄, 115 mM NaCl) beladen und anschließend 2 Std. lang in PBS-Puffer mit 3% w/v BSA (Sigma, Fraktion V) und 0,5% v/v Tween 20 (Sigma) blockiert. Diese "Antikörper-Fang-Membran" wurde zweimal mit PBS gewaschen, mit flüssigem LB-Medium getränkt, das 100 µg/ml Ampicillin und 1 mM IPTG (Isopropyl β-D-Thiogalactopyranosid, Biomol) enthielt, auf eine Agar-Platte mit der gleichen Medienzusammensetzung aufgebracht und mit der die E. coli-Kolonien tragenden ersten Membran bedeckt. Die Zellen auf diesem Filter-Stapel wurden über Nacht (14 Std.) bei Raumtemperatur inkubiert um eine Expression der Fv-Fragmente mit C-terminalem Peptidanhängsel zu erreichen. Die obere Membran wurde dann abgehoben und auf eine frische LB Agar Platte gelegt, die 100 µg/ml Ampicillin und 10 mg/ml Glucose enthielt, um die E. coli-Kolonien zur späteren Vermehrung bei 4°C aufzubewahren. Die "Fang"-Membran wurde vom Agar abgehoben und dreimal mit PBS/Tween (PBS mit 0,1% v/v Tween) gewaschen. Der Nachweis von immobilisierten Fv-Fragmenten, die Peptide mit Streptavidin-bindender Aktivität trugen, wurde durch einstündige Inkubation mit einem Streptavidin-alkalische Phosphatase-Konjugat erreicht, (1 : 2000 verdünnt in PBS/Tween), in Gegenwart von 2 µg/ml Avidin (Polyscience), das 10 Min. vorher zugefügt worden war. Nach gründlichem Waschen (3 × PBS/Tween; 2 × PBS) wurde die Membran in 10 ml AP-Puffer inkubiert (100 mM Tris pH 8,8; 100 mM NaCl; 5 mM MgCl₂), dem 30 µl BCIP-Stammlösung (50 mg/ml 5-Brom-4-Chlor-3-Indolyl-Phosphat 4-Toluidin Salz (Biomol) in Dimethyl-Formamid) und 5 µl NBT-Stammlösung (75 mg/ml Nitro-Blau Tetrazolium (Biomol) in 70% v/v Dimethyl Formamid) zugesetzt worden waren (Blake, M.S., Johnston, K.H. Russel-Jones, G.J. und Gotschlich, E.C. (1984) Anal. Biochem., 136, 175–179). Die chromogene Reaktion wurde nach 30–60 Min. durch mehrmaliges Waschen in destilliertem Wasser gestoppt und der Filter an der Luft getrocknet.

Fig. 3 zeigt jeweils 4 Kolonien von Escherichia coli TG1, transformiert mit den Plasmiden pASK46-p141H (1), pASK46-p111H (2), pASK46 (3) (als Negativkontrolle), und pASK46-p111L (4). Abgesehen von der Negativkon-

trolle wird in allen Fällen ein intensives Bindungssignal für das Streptavidin-Konjugat beobachtet. Ein etwas schwächeres Bindungssignal wurde mit den Plasmiden pASK46-p11XH, pASK46-p11XL und pASK46-p14XH nachgewiesen (nicht gezeigt).

5

Beispiel 2

SDS-PAGE und Western-Transfer

SDS-PAGE wurde in vertikalen Flachgelkammern unter Verwendung des Puffersystems von Fling, S.P. und Gregerson, D.S. (1986) Anal. Biochem., 155, 83–88) durchgeführt. Für den Western Transfer wurde das Gel nach der Elektrophorese in Transferpuffer (Elektrophorese-Laufpuffer mit 20% (v/v) Methanol) getränkt. Das Protein wurde mit der "Semi-dry"-Technik auf eine Immobilon-P Membran (Millipore) transferiert bei einer konstanten Stromstärke von 1 mA/cm² im Verlauf einer Stunde bei 4°C. Anschließend wurde die Membran mit PBS, das 0,5% v/v Tween und 3% BSA w/v enthielt, 1 Std. lang bei Raumtemperatur blockiert. Nach dreimaligem Waschen mit PBS/Tween wurde die Membran 10 Min. lang mit Avidin inkubiert (2 µg/ml in PBS/Tween), um das Biotincarboxyl-Carrierprotein zu komplexieren. Danach wurde mit einem Streptavidin-alkalische Phosphatase-Konjugat 1 : 4000 verdünnt in PBS/Tween eine Stunde lang inkubiert. Nach gründlichem Waschen (3 × PBS/Tween; 2 × PBS) wurde der Transfer unter Verwendung des BCIP/NBT Protokolls (siehe Beispiel 1) entwickelt.

Fig. 4 zeigt den Nachweis der entsprechenden Protein-Peptidfusionen im Western-Blot. Es wurden gleiche Aliquots der periplasmatischen Fraktionen von induzierten *Escherichia coli* TG1 Zellen, transformiert mit den Plasmiden pASK46-p141H (1), pASK46-p111H (2), pASK46-p111L (3) und pASK46 (4) (als Negativkontrolle) durch SDS-PAGE aufgetrennt. Nach dem Transfer der Proteine auf eine Immobilon-P Membran wurde die entsprechende Peptid-Protein Fusion mit dem Streptavidin-Konjugat spezifisch nachgewiesen. Ein etwas schwächeres Signal in einem analogen Experiment lieferten die Plasmide pASK46-p11XH, pASK46-p11XL und pASK46-p14XH (nicht gezeigt). Vergleichbare Resultate wurden erzielt, wenn das Gesamtzellprotein im SDS-PAGE aufgetrennt wurde.

Beispiel 3

30

ELISA

Sieben Reihen einer Mikrotiter-Platte mit 96 Vertiefungen (Falcon #3912) wurden mit 100 µl einer Lösung von 3 mg/ml Lysozym in 50 mM NaHCO₃, pH 9,6, bei 4°C über Nacht beladen. Anschließend wurde bei Raumtemperatur mit 2% w/v fettfreier Trockenmilch (BioRad) in PBS 2 Std. lang blockiert. Nach dem Waschen (3 × PBS/Tween) wurden je 50 µl der periplasmatischen Zellfraktionen der entsprechenden Klone als Verdünnungsreihe in PBS/Tween zupipettiert und 1 Std. lang inkubiert. Die Konzentrationen der Fv-Fragmente in den unverdünnten Fraktionen betrugen ungefähr 50–100 µg/ml (geschätzt durch SDS-PAGE). Nach dem Waschen (3 × PBS/Tween), wurden 50 µl Streptavidin-alkalische Phosphatase-Konjugat in einer Verdünnung von 1 : 1000 in PBS/Tween zugesetzt und 1 Std. lang inkubiert. Ungebundenes Konjugat wurde durch jeweils zweifaches Waschen mit PBS/Tween und PBS entfernt und 100 µl NPP-Lösung (0,5 mg/ml p-Nitrophenyl-Phosphat; 0,9 M Diethanol-Amin, pH 9,6; 1 mM MgCl₂) zupipettiert. Das Farbsignal wurde 5–10 Min. lang entwickelt und durch Zugabe von 100 µl 10 mM EDTA, pH 8,0, gestoppt. Die Absorptionswerte wurden unter Verwendung eines Mikrotiterplatten-Photometers bestimmt, und die Daten wurden als A₄₀₅ – A₄₅₀ Differenzwerte nach Abzug des Blindwertes für jede Verdünnung aufgezeichnet. Die Blindwerte wurden durch eine Verdünnungsreihe der periplasmatischen Zellfraktion bestimmt, die das D1.3 Fv-Fragment ohne Peptidanhängsel enthielt, und erwiesen sich als konstant über den gesamten Verdünnungsbereich (A₄₀₅ – A₄₅₀ = 0,199 ± 0,009).

Fig. 5 zeigt die beobachteten Bindungssignale für das Streptavidin-Konjugat mit Verdünnungsreihen der Periplasmafraktionen von induzierten *Escherichia coli* TG1-Zellen, transformiert mit den entsprechenden Plasmiden.

50

Beispiel 4

Proteinreinigung durch Streptavidin-Affinitätschromatographie

Eine Säule, die mit 6 ml Streptavidin-Agarose (ungefähr 1 mg Streptavidin pro 1 ml Gel) gepackt war, wurde mit 10 Volumen von 50 mM Tris, pH 8,0 equilibriert. Die periplasmatische Zellfraktion von 0,5 l der mit dem entsprechenden Plasmid transformierten *E. coli*-Zellen wurde auf die Säule aufgetragen und mit dem Tris-Puffer nachgewaschen. Das hybride D1.3 Fv-Fragment wurde dann mit einer Lösung von 1 mM Iminobiotin (Sigma) oder 5 mM Liponsäure (Sigma) in dem gleichen Puffer spezifisch eluiert. Da diese Biotin-Analoga sehr viel schwächer an Streptavidin binden als Biotin selbst (Grün, N.M. (1975) Adv. Protein Chem., 29, 85–133), konnte die Säule durch einfaches Waschen mit 10 Volumen Tris-Puffer regeneriert werden. Alle chromatographischen Schritte wurden bei einer Durchflußrate von 30 ml/Std. bei einer Temperatur von 4°C durchgeführt. Die Streptavidin-Affinitätschromatographie des löslichen Teils des gesamten Zellproteins wurde bei einer Durchflußrate von ungefähr 20 ml/Std. unter sonst identischen Bedingungen durchgeführt. Die Ausbeute an gereinigtem hybriden Fv-Fragment lag normalerweise im Bereich von 0,5 mg/L · OD₅₅₀ *E. coli*-Kultur. Zur Streptavidin-Affinitätsreinigung des D1.3 Fv-Lysozym-Komplexes wurde eine periplasmatische Zellfraktion, die das D1.3 Fv(p111L) Fragment enthielt und gegen 50 mM NaH₂PO₄, pH 7,0, 115 mM NaCl, 1 mM EDTA dialysiert worden war, mit der ca. dreifachen molaren Menge an Lysozym versetzt. Nach 1-stündiger Inkubation bei 4°C und

anschließender Zentrifugation wurde die resultierende Proteinlösung der Streptavidin-Agarose Affinitätschromatographie unterworfen (wie oben), wobei der letztgenannte Puffer verwendet wurde.

Fig. 6A zeigt die Elutionsprofile (Absorption bei 280 nm) der Streptavidin-Affinitätschromatographie, welche mit den Periplasmafraktionen von *Escherichia coli* TG1-Zellen, die mit pASK46-p111H bzw. pASK46-p111L transformiert waren, durchgeführt wurde. Darunter sind mit Coomassie-Brilliantblau (Serva) angefärbte SDS-Polyacrylamidgels der nach Anlegen einer Iminobiotin-Lösung erhaltenen Fraktionen wiedergegeben. Daraus ist zu ersehen, daß beide Untereinheiten des Fv-Fragments gemeinsam in reiner Form und spezifisch eluiert wurden, obwohl nur jeweils eine Kette des heterodimeren Proteins mit dem Streptavidin-Affinitätspeptid fusioniert war. Im Fall von pASK46-p111L wiesen die beiden Ketten gleiche Mobilität in der SDS-PAGE auf. Die Funktionalität des gereinigten Proteins wurde daher durch ein ELISA-Experiment (wie in Beispiel 3) zusätzlich nachgewiesen. Bei einer analog durchgeführten Chromatographie, bei der das Gesamtzellprotein von *Escherichia coli* TG1-Zellen, die mit pASK46-p111H transformiert waren, aufgetragen wurde, wurde ebenfalls das rekombinante Protein als intaktes Heterodimer in reiner Form erhalten.

Fig. 6B stellt eine analoge Streptavidin-Affinitätschromatographie dar, wobei hier Lysozym, das Antigen des D1.3 Fv-Fragments, der Periplasmafraktion von *Escherichia coli* TG1-Zellen, transformiert mit pASK46-p111L, zugesetzt worden war. Das unter dem Elutionsprofil wiedergegebene, mit Coomassie-Brilliantblau (Serva) angefärbte SDS-Polyacrylamidgel zeigt die durch Elution mit der Iminobiotin-Lösung erhaltene Produktfraktion (2) und zum Vergleich gereinigtes Lysozym (3) sowie das separat gereinigte rekombinante Fv-Fragment (4). Spur (1) gibt den Molekulargewichtsstandard wieder. Es wird ersichtlich, daß der Immunkomplex aus rekombinantem Fv-Fragment und dem Antigen Lysozym spezifisch gereinigt worden ist.

Beispiel 5

Konstruktion und Verwendung von pASK60-Strep

pASK60-Strep wurde ausgehend von pASK40 (Skerra, A., Pfitzinger, I. und Plückthun, A. (1991) *Bio/Technology*, 9, 273–278) unter Verwendung ortsgerichteter Mutagenese und PCR hergestellt. Eine Restriktionskarte sowie die gesamte DNA-Sequenz von pASK60-Strep sind in den Fig. 7 und 8 gezeigt. Der Polylinker auf pASK60-Strep enthält einen verbesserten Satz singulärer Restriktionsschnittstellen, einschließlich zweier Restriktionsschnittstellen, die direkt am 3'-Ende der für das OmpA-Signalpeptid kodierenden Region liegen, und wird von einer für das Streptavidin-Bindungspeptid (Ser–Ala–)Trp–Arg–His–Pro–Gln–Phe–Gly–Gly kodierenden DNA-Sequenz gefolgt (vgl. Fig. 9). Die direkte Expression eines Fremdgens in *E. coli*, d. h. ohne Verwendung der OmpA-Signalsequenz, kann erreicht werden unter Verwendung der XbaI-Schnittstelle und Rekonstruktion des 16 Basenpaare umfassenden Bereichs zwischen dem Stop-Codon des lacZ Mini-Cistrons und dem Start-Codon des Strukturgens (vormals OmpA-Signalsequenz). Um dagegen eine Fusion mit der OmpA-Signalsequenz herzustellen, die auf pASK60-Strep kodiert ist, kann das Strukturgen direkt unter Verwendung der StuI- oder BsaI-Schnittstellen einkloniert werden. StuI (Erkennungssequenz "AGG CCT") erzeugt ein glattes Ende nach dem ersten Nukleotid des letzten Codons (Ala) der Signal-Sequenz. Eine präzise Fusion kann somit unter Zuhilfenahme einer kompatiblen Restriktionsstelle (z. B. StuI, PvuII, NruI) direkt vor dem Strukturgen geschaffen werden. BsaI ist ein Restriktionsenzym vom Typ IIa, dessen Schneidestelle von der Erkennungsstelle entfernt liegt ("GAGACC", unterstrichen in Fig. 9). Durch Schneiden mit diesem Enzym wird ein 5'-überhängendes DNA-Ende erzeugt ("GGCC", in kleinen Buchstaben gedruckt in Fig. 9), das am äußersten 3'-Ende der Signal-Sequenz liegt, ohne in die Kodierungsregion des muren Teiles des zu klonierenden Gens hineinzuragen. Dieses kohäsive Ende kann mit den Enden, die von den Restriktionsenzymen EaeI und EagI erzeugt werden, ligiert werden. Weiterhin ist es möglich, z. B. durch PCR bei der Herstellung des zu klonierenden DNA-Fragments, eine kompatible Restriktionsstelle einzuführen, die die reife aminoterminal kodierende Region des Gens nicht beeinträchtigt, wenn die BsaI-(oder ein anderes IIa-Enzym, wie BspMI etc.) Erkennungssequenz auf die gegenüberliegende Seite des kohäsiven Endes, das erzeugt werden soll, plaziert wird. Die Fusion des Strukturgens an die für das Streptavidin-Bindungspeptid kodierende Region kann unter Verwendung der Eco47III-Restriktionsschnittstelle (Erkennungssequenz "AGC GCT") erreicht werden, was zu einem glatten Schnitt direkt zwischen dem Ser-Codon ("AGC", in kleinen Buchstaben gedruckt in Fig. 9) und dem Ala-Codon vor der eigentlichen Peptidsequenz nach Anspruch 1 führt. Wenn das Ser-Codon wiederhergestellt werden soll, können verschiedene Restriktionsstellen für die Erzeugung eines kompatiblen Endes benutzt werden, außer Eco47III z. B. auch Scal und NruI.

Beispiel 6

Expression und Nachweis einer löslichen Domäne des LDL-Rezeptors mit pASK60-Strep

Der menschliche "Low density lipoprotein"-Rezeptor besteht, abgesehen von einer N-terminalen Signalsequenz, die im muren Rezeptor abgespalten ist, aus fünf Proteindomänen, von denen die vierte die Transmembrananteile darstellt. Zur Klonierung der ersten, N-terminalen Proteindomäne, die die Aminosäuren 1 bis 292 umfaßt, wurde das Plasmid pLDLR3 (ATCC Nr. 57004) (Yamamoto, T., Davis, C.G., Brown, M.S., Schneider, W.J., Casey, M.L., Goldstein, J.L. und Russell, D.W. (1984) *Cell* 39, 27–38) verwendet. Das entsprechende DNA-Fragment wurde mit Hilfe der Oligodesoxynukleotid-Primer "5'-TAG CAA CGG CCG CAG TGG GCG ACA GAT GT" (für den N-Terminus, mit der Restriktionsschnittstelle EagI) und "5'-TTC GTT AGT ACT GCA CTC TTT GAT GGG TTC" (für den C-Terminus, mit der Restriktionsschnittstelle Scal) durch PCR amplifiziert, isoliert und mit den Restriktionsenzymen EagI und Scal geschnitten. Dieses DNA-Fragment wurde in pASK60-Strep

über dessen Restriktionsschnittstellen BsaI und Eco47III unter Verwendung des E.coli-Stamms K12 JM83 kloniert. Nach Überprüfung des erhaltenen Plasmids durch Restriktionsanalyse und DNA-Sequenzierung wurde von einem Klon eine Kultur hergestellt und die Proteinexpression induziert, wie weiter ben beschrieben. Das Gesamtzellprotein der erhaltenen Zellen wurde in einer 15%igen SDS-PAGE aufgetrennt und durch Western-
 5 Transfer auf eine Immobilon P-Membran überführt. Mit Hilfe des Streptavidin-alkalische Phosphatase-Konjugats (vgl. Beispiel 2) konnte die rekombinante Rezeptordomäne von ca. 40 000 Dalton spezifisch nachgewiesen werden.

Patentansprüche

- 10 1. Peptid, umfassend die Aminosäuresequenz Trp—X—His—Pro—Gln—Phe—Y—Z, wobei X eine beliebige Aminosäure darstellt und Y und Z entweder beide Gly oder Y Glu und Z Arg oder Lys bedeuten.
2. Fusionsprotein, dadurch gekennzeichnet, daß es aus der Aminosäuresequenz eines vollständigen Proteins, einer Protein-Mutante, wie einer Deletions- oder Substitutions-Mutante, oder einem Proteinteil verbunden
 15 mit der Sequenz des Peptids nach Anspruch 1 besteht.
3. Expressionsvektor, dadurch gekennzeichnet, daß er, exprimierbar unter der Kontrolle eines geeigneten Promotors und gegebenenfalls Operators, eine DNA-Sequenz enthält, welche für ein Peptid nach Anspruch 1 kodiert sowie eine Restriktionsschnittstelle in 5'- oder/und 3'-Richtung anschließend an diese DNA-Sequenz aufweist, welche die Einbringung einer weiteren DNA-Sequenz erlaubt, die für ein zu exprimierendes
 20 Protein oder einen Proteinteil kodiert.
4. Verfahren zur Herstellung eines rekombinanten Proteins durch Expression einer dafür kodierenden DNA-Sequenz in geeigneten Wirtszellen nach an sich bekannten Methoden, dadurch gekennzeichnet, daß man eine DNA-Sequenz verwendet, welche für ein Fusionsprotein gemäß Anspruch 2 kodiert und gegebenenfalls die Anwesenheit des Expressionsproduktes über ein Konjugat von Streptavidin und einer Markie-
 25 rung nachweist bzw. die Abtrennung des gewünschten Proteins als Fusionsprotein über eine Streptavidinaffinitätschromatographie bewirkt.
5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß man ein Konjugat aus Streptavidin und einer Enzymmarkierung verwendet.
6. Verfahren nach Anspruch 5, dadurch gekennzeichnet, daß man als Enzymmarkierung alkalische Phosphatase verwendet.
 30
7. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß man zur Affinitätschromatographie eine Streptavidin-Agarosematrix verwendet.
8. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß man bei der Affinitätschromatographie das Fusionsprotein durch Zugabe von natürlichen Liganden des Streptavidins, insbesondere Biotin, syntheti-
 35 schen Derivaten davon, insbesondere 2-Iminobiotin oder Liponsäure, oder synthetischen Peptiden nach Anspruch 1 eluiert.

Hierzu 30 Seite(n) Zeichnungen

Figur 1

DNS-Sequenz des Plasmids pASK46

Gezeigt sind Strang (mit kodierten Aminosäuresequenzen in allen drei Leserahmen) und Gegenstrang mit singulären Restriktionsschnittstellen.

```

      GCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCA
1  -----+-----+-----+-----+-----+-----+-----+ 60
      CGCGGGTTATGCGTTTGGCGGAGAGGGGCGCGCAACCGGCTAAGTAATTACGTCGACCGT

a:      AlaProAsnThrGlnThrAlaSerProArgAlaLeuAlaAspSerLeuMetGlnLeuAla  -
b:      ArgProIleArgLysProProLeuProAlaArgTrpProIleHisEndCysSerTrpHis  -
c:      AlaGlnTyrAlaAsnArgLeuSerProArgValGlyArgPheIleAsnAlaAlaGlyThr  -

      CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCT
61 -----+-----+-----+-----+-----+-----+-----+ 120
      GCTGTCCAAAGGGCTGACCTTTTCGCCCCGTCACTCGCGTTGCGTTAATTACACTCAATCGA

a:      ArgGlnValSerArgLeuGluSerGlyGlnEndAlaGlnArgAsnEndCysGluLeuAla  -
b:      AspArgPheProAspTrpLysAlaGlySerGluArgAsnAlaIleAsnValSerEndLeu  -
c:      ThrGlyPheProThrGlyLysArgAlaValSerAlaThrGlnLeuMetEndValSerSer  -

      CACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAAT
121 -----+-----+-----+-----+-----+-----+-----+ 180
      GTGAGTAATCCGTGGGGTCCGAAATGTGAAATACGAAGGCCGAGCATACAACACACCTTA

a:      HisSerLeuGlyThrProGlyPheThrLeuTyrAlaSerGlySerTyrValValTrpAsn  -
b:      ThrHisEndAlaProGlnAlaLeuHisPheMetLeuProAlaArgMetLeuCysGlyIle  -
c:      LeuIleArgHisProArgLeuTyrThrLeuCysPheArgLeuValCysCysValGluLeu  -

                                     H
                                     i
                                     n
                                     d
                                     I
                                     I
                                     I

      TGTGAGCGGATAACAATTTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGC
181 -----+-----+-----+-----+-----+-----+-----+ 240
      ACACTCGCCTATTGTTAAAGTGTGTCCTTTGTCGATACTGGTACTAATGCGGTTTCAACG

a:      CysGluArgIleThrIleSerHisArgLysGlnLeuEndProEndLeuArgGlnAlaCys  -
b:      ValSerGlyEndGlnPheHisThrGlyAsnSerTyrAspHisAspTyrAlaLysLeuAla  -
c:      EndAlaAspAsnAsnPheThrGlnGluThrAlaMetThrMetIleThrProSerLeuHis  -

      ATGCAAATTCCTATTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGC
241 -----+-----+-----+-----+-----+-----+-----+ 300
      TACGTTTAAGATAAAGTTCCTCTGTCTAGTATTACTTTATGGATAACGGATGCCGTGGCG

a:      MetGlnIleLeuPheGlnGlyAspSerHisAsnGluIleProIleAlaTyrGlySerArg  -
b:      CysLysPheTyrPheLysGluThrValIleMetLysTyrLeuLeuProThrAlaAlaAla  -
c:      AlaAsnSerIleSerArgArgGlnSerEndEndAsnThrTyrCysLeuArgGlnProLeu  -

```

E
C
O
OP
P lp
s Ou
t 9M
I II
/

301 TGGATTGTTATTACTCGCTGCCCAACCAGCGATGGCCCAGGTGCAGCTGCAGGAGTCAGG 360
-----+-----+-----+-----+-----+-----+-----+-----+
ACCTAACAATAATGAGCGACGGGTTGGTTCGCTACCGGGTCCACGTCGACGTCCTCAGTCC

a: TrpIleValIleThrArgCysProThrSerAspGlyProGlyAlaAlaAlaGlyValArg -
b: GlyLeuLeuLeuLeuAlaAlaGlnProAlaMetAlaGlnValGlnLeuGlnGluSerGly -
c: AspCysTyrTyrSerLeuProAsnGlnArgTrpProArgCysSerCysArgSerGlnAsp -

N
a
r
I

361 ACCTGGCCTGGTGGCGCCCTCACAGAGCCTGTCCATCACATGCACCGTCTCAGGGTTCTC 420
-----+-----+-----+-----+-----+-----+-----+-----+
TGGACCGGACCACCGCGGGAGTGTCTCGGACAGGTAGTGTACGTGGCAGAGTCCCAAGAG

a: ThrTrpProGlyGlyAlaLeuThrGluProValHisHisMetHisArgLeuArgValLeu -
b: ProGlyLeuValAlaProSerGlnSerLeuSerIleThrCysThrValSerGlyPheSer -
c: LeuAlaTrpTrpArgProHisArgAlaCysProSerHisAlaProSerGlnGlySerHis -

421 ATTAACCGCTATGGTGTAACTGGGTTTCGCCAGCCTCCAGGAAAGGTCTGGAGTGGCT 480
-----+-----+-----+-----+-----+-----+-----+-----+
TAATTGGCCGATACCACATTTGACCCAAGCGGTCCGAGGTCTTTCCAGACCTCACCAG

a: IleAsnArgLeuTrpCysLysLeuGlySerProAlaSerArgLysGlySerGlyValAla -
b: LeuThrGlyTyrGlyValAsnTrpValArgGlnProProGlyLysGlyLeuGluTrpLeu -
c: EndProAlaMetValEndThrGlyPheAlaSerLeuGlnGluArgValTrpSerGlyTrp -

481 GGGAAATGATTTGGGGTGATGGAAACACAGACTATAATTTCAGCTCTCAAATCCAGACTGAG 540
-----+-----+-----+-----+-----+-----+-----+-----+
CCCTTACTAAACCCCACTACCTTTGTGTCTGATATTAAGTCGAGAGTTTAGGTCTGACTC

a: GlyAsnAspLeuGlyEndTrpLysHisArgLeuEndPheSerSerGlnIleGlnThrGlu -
b: GlyMetIleTrpGlyAspGlyAsnThrAspTyrAsnSerAlaLeuLysSerArgLeuSer -
c: GluEndPheGlyValMetGluThrGlnThrIleIleGlnLeuSerAsnProAspEndAla -

541 CATCAGCAAGGACAACCTCCAAGAGCCAAGTTTTCTTAAAAATGAACAGTCTGCACACTGA 600
-----+-----+-----+-----+-----+-----+-----+-----+
GTAGTCGTTCTTGTGAGGTTCTCGGTTCAAAGAATTTTTACTTGTTCAGACGTGTGACT

a: HisGlnGlnGlyGlnLeuGlnGluProSerPheLeuLysAsnGluGlnSerAlaHisEnd -
b: IleSerLysAspAsnSerLysSerGlnValPheLeuLysMetAsnSerLeuHisThrAsp -
c: SerAlaArgThrThrProArgAlaLysPheSerEndLysEndThrValCysThrLeuMet -

S
t
y
I

601 TGACACAGCCAGGTACTACTGTGCCAGAGAGAGATTATAGGCTTGACTACTGGGGCCA
 -----+-----+-----+-----+-----+ 660
 ACTGTGTCCGGTCCATGATGACACGGTCTCTCTCTCTAATATCCGAACTGATGACCCCGGT

a: EndHisSerGlnValLeuLeuCysGlnArgGluArgLeuEndAlaEndLeuLeuGlyPro -
 b: AspThrAlaArgTyrTyrCysAlaArgGluArgAspTyrArgLeuAspTyrTrpGlyGln -
 c: ThrGlnProGlyThrThrValProGluArgGluIleIleGlyLeuThrThrGlyAlaLys -

B
S
D t
s E
a I
I I

661 AGGCACCACGGTCACCGTCTCCTCATAATAAGAGCTATGGGAGCTTGACATGCAAATTCTA
 -----+-----+-----+-----+-----+ 720
 TCCGTGGTGGCAGTGGCAGAGAGTATTATTCTCGATACCCTCGAACGTACGTTTAAGAT

a: ArgHisHisGlyHisArgLeuLeuIleIleArgAlaMetGlyAlaCysMetGlnIleLeu -
 b: GlyThrThrValThrValSerSerEndEndGluLeuTrpGluLeuAlaCysLysPheTyr -
 c: AlaProArgSerProSerProHisAsnLysSerTyrGlySerLeuHisAlaAsnSerIle -

721 TTTCAAGGAGACAGTCATAATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATT
 -----+-----+-----+-----+-----+ 780
 AAAGTTCCTCTGTCTAGTATTACTTTATGGATAACGGATGCCGTCGGCGACCTAACAATAA

a: PheGlnGlyAspSerHisAsnGluIleProIleAlaTyrGlySerArgTrpIleValIle -
 b: PheLysGluThrValIleMetLysTyrLeuLeuProThrAlaAlaAlaGlyLeuLeuLeu -
 c: SerArgArgGlnSerEndEndAsnThrTyrCysLeuArgGlnProLeuAspCysTyrTyr -

S
s
t
I

R
l
e
A
I

781 ACTCGCTGCCCAACCAGCGATGGCCGACATCGAGCTCACCCAGTCTCCAGCCTCCCTTTC
 -----+-----+-----+-----+-----+ 840
 TGAGCGACGGGTGGTTCGCTACCGGCTGTAGCTCGAGTGGGTGAGAGGTCGGAGGGAAG

a: ThrArgCysProThrSerAspGlyArgHisArgAlaHisProValSerSerLeuProPhe -
 b: LeuAlaAlaGlnProAlaMetAlaAspIleGluLeuThrGlnSerProAlaSerLeuSer -
 c: SerLeuProAsnGlnArgTrpProThrSerSerSerProSerLeuGlnProPropheLeu -

841 TCGCTCTGTGGGAGAACTGTCAACATCACATGTCGAGCAAGTGGGAATATTCACAATTA
 -----+-----+-----+-----+-----+ 900
 ACGCAGACACCCTCTTTGACAGTGGTAGTGACAGCTCGTTACCCCTTATAAGTGTTAAT

a: CysValCysGlyArgAsnCysHisHisHisMetSerSerLysTrpGluTyrSerGlnLeu -
 b: AlaSerValGlyGluThrValThrIleThrCysArgAlaSerGlyAsnIleHisAsnTyr -
 c: ArgLeuTrpGluLysLeuSerProSerHisValGluGlnValGlyIlePheThrIleIle -

```

          TTTAGCATGGTATCAGCAGAAACAGGGAAAATCTCCTCAGCTCCTGGTCTATTATACAAC
901  -----+-----+-----+-----+-----+-----+-----+-----+ 960
          AAATCGTACCATAGTCGTCTTTGTCCCTTTTAGAGGAGTCGAGGACCAGATAATATGTTG

a:      PheSerMetValSerAlaGluThrGlyLysIleSerSerAlaProGlyLeuLeuTyrAsn -
b:      LeuAlaTrpTyrGlnGlnLysGlnGlyLysSerProGlnLeuLeuValTyrTyrThrThr -
c:      EndHisGlyIleSerArgAsnArgGluAsnLeuLeuSerSerTrpSerIleIleGlnGln -

          AACCTTAGCAGATGGTGTGCCATCAAGGTTTCAGTGGCAGTGGATCAGGAACACAATATTC
961  -----+-----+-----+-----+-----+-----+-----+-----+ 1020
          TTGGAATCGTCTACCACACGGTAGTTCCAAGTCACCGTCACCTAGTCCTTGTGTTATAAG

a:      AsnLeuSerArgTrpCysAlaIleLysValGlnTrpGlnTrpIleArgAsnThrIlePhe -
b:      ThrLeuAlaAspGlyValProSerArgPheSerGlySerGlySerGlyThrGlnTyrSer -
c:      ProEndGlnMetValCysHisGlnGlySerValAlaValAspGlnGluHisAsnIleLeu -

                                     H
                                     i
                                     n
                                     c
                                     I
                                     I
          TCTCAAGATCAACAGCCTGCAACCTGAAGATTTTGGGAGTTATTACTGTCAACATTTTGT
1021 -----+-----+-----+-----+-----+-----+-----+-----+ 1080
          AGAGTTCTAGTTGTTCGGACGTTGGACTTCTAAACCCTCAATAATGACAGTTGTAAAAAC

a:      SerGlnAspGlnGlnProAlaThrEndArgPheTrpGluLeuLeuLeuSerThrPheLeu -
b:      LeuLysIleAsnSerLeuGlnProGluAspPheGlySerTyrTyrCysGlnHisPheTrp -
c:      SerArgSerThrAlaCysAsnLeuLysIleLeuGlyValIleThrValAsnIlePheGly -

                                     AX
                                     vh
                                     ao
                                     II
                                     /
          GAGTACTCCTCGGACGTTTCGGTGGAGGCACCAAGCTCGAGATCAAACGGGAACAAAACT
1081 -----+-----+-----+-----+-----+-----+-----+-----+ 1140
          CTCATGAGGAGCCTGCAAGCCACCTCCGTGGTTCGAGCTCTAGTTTGCCCTTGTTTTGA

a:      GluTyrSerSerAspValArgTrpArgHisGlnAlaArgAspGlnThrGlyThrLysThr -
b:      SerThrProArgThrPheGlyGlyGlyThrLysLeuGluIleLysArgGluGlnLysLeu -
c:      ValLeuLeuGlyArgSerValGluAlaProSerSerArgSerAsnGlyAsnLysAsnSer -

                                     B
                                     C
                                     l
                                     I
          CATCTCAGAAGAGGATCTGAATTAATAATGATCAAACGGTAATAAGGATCAGCTTGACCT
1141 -----+-----+-----+-----+-----+-----+-----+-----+ 1200
          GTAGAGTCTTCTCCTAGACTTAATTATTACTAGTTTGCCATTATTCCTAGTCGAACCTGGA

a:      HisLeuArgArgGlySerGluLeuIleMetIleLysArgEndEndGlySerAlaEndPro -
b:      IleSerGluGluAspLeuAsnEndEndEndSerAsnGlyAsnLysAspGlnLeuAspLeu -
c:      SerGlnLysArgIleEndIleAsnAsnAspGlnThrValIleArgIleSerLeuThrCys -

```

A
P
a
B
I

GTGAAGTGAAAAATGGCGCACATTGTGCGACATTTTTTTTGTCTGCCGTTTACCGCTACT
 1201 -----+-----+-----+-----+-----+-----+ 1260
 CACTTCACTTTTACCGCGTGTAACACGCTGTAAAAAAACAGACGGCAATGGCGATGA

a: ValLysEndLysMetAlaHisIleValArgHisPhePheCysLeuProPheThrAlaThr -
 b: EndSerGluLysTrpArgThrLeuCysAspIlePhePheValCysArgLeuProLeuLeu -
 c: GluValLysAsnGlyAlaHisCysAlaThrPhePheLeuSerAlaValTyrArgTyrCys -

B
a
m
H
I

GCGTCACGGATCCCCACGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGTGTGGTGGTTA
 1261 -----+-----+-----+-----+-----+ 1320
 CGCAGTGCCTAGGGGTGCGCGGGACATCGCCGCGTAATTCGCGCCGCCACACCAAT

a: AlaSerArgIleProThrArgProValAlaAlaHisEndAlaArgArgValTrpTrpLeu -
 b: ArgHisGlySerProArgAlaLeuEndArgArgIleLysArgGlyGlyCysGlyGlyTyr -
 c: ValThrAspProHisAlaProCysSerGlyAlaLeuSerAlaAlaGlyValValValThr -

CGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCC
 1321 -----+-----+-----+-----+-----+ 1380
 GCGCGTCGCACTGGCGATGTGAACGGTTCGCGGGATCGCGGGCGAGGAAAGCGAAAGAAGG

a: ArgAlaAlaEndProLeuHisLeuProAlaProEndArgProLeuLeuSerLeuSerSer -
 b: AlaGlnArgAspArgTyrThrCysGlnArgProSerAlaArgSerPheArgPheLeuPro -
 c: ArgSerValThrAlaThrLeuAlaSerAlaLeuAlaProAlaProPheAlaPhePhePro -

N
a
e
I

CTTCCTTTCTCGCCACGTTTCGCGGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTT
 1381 -----+-----+-----+-----+-----+ 1440
 GAAGGAAAGAGCGGTGCAAGCGGCCGAAAGGGGCGAGTTTCGAGATTTAGCCCCGAGGGAA

a: LeuProPheSerProArgSerProAlaPheProValLysLeuEndIleGlyGlySerLeu -
 b: PheLeuSerArgHisValArgArgLeuSerProSerSerSerLysSerGlyAlaProPhe -
 c: SerPheLeuAlaThrPheAlaGlyPheProArgGlnAlaLeuAsnArgGlyLeuProLeu -

TAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATG
 1441 -----+-----+-----+-----+-----+ 1500
 ATCCCAAGGCTAAATCACGAAATGCCGTGGAGCTGGGGTTTTTTGAACTAATCCCACTAC

a: EndGlySerAspLeuValLeuTyrGlyThrSerThrProLysAsnLeuIleArgValMet -
 b: ArgValProIleEndCysPheThrAlaProArgProGlnLysThrEndLeuGlyEndTrp -
 c: GlyPheArgPheSerAlaLeuArgHisLeuAspProLysLysLeuAspEndGlyAspGly -

B
S
a
A
I

1501 GTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1560
 CAAGTGCATCACCCGGTAGCGGGACTATCTGCCAAAAGCGGGAACTGCAACCTCAGGT

a: ValHisValValGlyHisArgProAspArgArgPhePheAlaLeuEndArgTrpSerPro -
 b: PheThrEndTrpAlaIleAlaLeuIleAspGlyPheSerProPheAspValGlyValHis -
 c: SerArgSerGlyProSerProEndEndThrValPheArgProLeuThrLeuGluSerThr -

1561 CGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1620
 GCAAGAAATTATCACCTGAGAACAAGGTTTGACCTTGTTGTGAGTTGGGATAGAGCCAGA

a: ArgSerLeuIleValAspSerCysSerLysLeuGluGlnHisSerThrLeuSerArgSer -
 b: ValLeuEndEndTrpThrLeuValProAsnTrpAsnAsnThrGlnProTyrLeuGlyLeu -
 c: PhePheAsnSerGlyLeuLeuPheGlnThrGlyThrThrLeuAsnProIleSerValTyr -

1621 ATTCTTTTGATTATATAAGGGATTTGCGGATTTTCGGCCTATTGGTTAAAAAATGAGCTGA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1680
 TAAGAAAACATAATATCCCTAAAACGGCTAAAGCCGGATAACCAATTTTTTACTCGACT

a: IleLeuLeuIleTyrLysGlyPheCysArgPheArgProIleGlyEndLysMetSerEnd -
 b: PhePheEndPheIleArgAspPheAlaAspPheGlyLeuLeuValLysLysEndAlaAsp -
 c: SerPheAspLeuEndGlyIleLeuProIleSerAlaTyrTrpLeuLysAsnGluLeuIle -

1681 TTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTTCAGGTGGCA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1740
 AAATTGTTTTTAAATTGCGCTTAAATTTGTTTTATAATTGCAAATGTTAAAGTCCACCGT

a: PheAsnLysAsnLeuThrArgIleLeuThrLysTyrEndArgLeuGlnPheGlnValAla -
 b: LeuThrLysIleEndArgGluPheEndGlnAsnIleAsnValTyrAsnPheArgTrpHis -
 c: EndGlnLysPheAsnAlaAsnPheAsnLysIleLeuThrPheThrIleSerGlyGlyThr -

1741 CTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTTATTTTCTAAATACATTCAAATA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1800
 GAAAAGCCCCCTTACACGCGCCTTGGGGATAACAAATAAAAAGATTTATGTAAGTTTAT

a: LeuPheGlyGluMetCysAlaGluProLeuPheValTyrPheSerLysTyrIleGlnIle -
 b: PheSerGlyLysCysAlaArgAsnProTyrLeuPheIlePheLeuAsnThrPheLysTyr -
 c: PheArgGlyAsnValArgGlyThrProIleCysLeuPhePheEndIleHisSerAsnMet -

1801 TGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 1860
 ACATAGGCGAGTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCTTCT

a: CysIleArgSerEndAspAsnAsnProAspLysCysPheAsnAsnIleGluLysGlyArg -
 b: ValSerAlaHisGluThrIleThrLeuIleAsnAlaSerIleIleLeuLysLysGluGlu -
 c: TyrProLeuMetArgGlnEndProEndEndMetLeuGlnEndTyrEndLysArgLysSer -

1861 GTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTGC GG CATTTTGCCTTC
 -----+-----+-----+-----+-----+-----+ 1920
 CATACTCATAAGTTGTAAAGGCACAGCGGGAATAAGGGAACCGCGTAAACGGAAG
 a: ValEndValPheAsnIleSerValSerProLeuPheProPheLeuArgHisPheAlaPhe -
 b: TyrGluTyrSerThrPheProCysArgProTyrSerLeuPheCysGlyIleLeuProSer -
 c: MetSerIleGlnHisPheArgValAlaLeuIleProPhePheAlaAlaPheCysLeuPro -

1921 CTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTG
 -----+-----+-----+-----+-----+-----+ 1980
 GACAAAACGAGTGGGTCTTTGCGACCACTTTTCATTTTCTACGACTTCTAGTCAACCCAC
 a: LeuPheLeuLeuThrGlnLysArgTrpEndLysEndLysMetLeuLysIleSerTrpVal -
 b: CysPheCysSerProArgAsnAlaGlyGluSerLysArgCysEndArgSerValGlyCys -
 c: ValPheAlaHisProGluThrLeuValLysValLysAspAlaGluAspGlnLeuGlyAla -

1981 CACGAGTGGGTTACATCGAAGTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCC
 -----+-----+-----+-----+-----+-----+ 2040
 GTGCTCACCAATGTAGCTTGACCTAGAGTTGTCGCCATTCTAGGAACCTCAAAAGCGG
 a: HisGluTrpValThrSerAsnTrpIleSerThrAlaValArgSerLeuArgValPheAla -
 b: ThrSerGlyLeuHisArgThrGlySerGlnGlnArgEndAspProEndGluPheSerPro -
 c: ArgValGlyTyrIleGluLeuAspLeuAsnSerGlyLysIleLeuGluSerPheArgPro -

X
 m
 n
 I

2041 CCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTAT
 -----+-----+-----+-----+-----+-----+ 2100
 GGCTTCTTGCAAAAGGTTACTACTCGTGAAAATTTCAAGACGATACACCGGCCATAATA
 a: ProLysAsnValPheGlnEndEndAlaLeuLeuLysPheCysTyrValAlaArgTyrTyr -
 b: ArgArgThrPheSerAsnAspGluHisPheEndSerSerAlaMetTrpArgGlyIleIle -
 c: GluGluArgPheProMetMetSerThrPheLysValLeuLeuCysGlyAlaValLeuSer -

2101 CCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCCGATACACTATTCTCAGAATGACT
 -----+-----+-----+-----+-----+-----+ 2160
 GGGCATAACTGCGGCCCGTTCTCGTTGAGCCAGCGCGTATGTGATAAGAGTCTTACTGA
 a: ProValLeuThrProGlyLysSerAsnSerValAlaAlaTyrThrIleLeuArgMetThr -
 b: ProTyrEndArgArgAlaArgAlaThrArgSerProHisThrLeuPheSerGluEndLeu -
 c: ArgIleAspAlaGlyGlnGluGlnLeuGlyArgArgIleHisTyrSerGlnAsnAspLeu -

2161 TGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAAT
 -----+-----+-----+-----+-----+-----+ 2220
 ACCAACTCATGAGTGGTCAGTGTCTTTTCGTAGAATGCCTACCGTACTGTCTTCTCTTA
 a: TrpLeuSerThrHisGlnSerGlnLysSerIleLeuArgMetAlaEndGlnEndGluAsn -
 b: GlyEndValLeuThrSerHisArgLysAlaSerTyrGlyTrpHisAspSerLysArgIle -
 c: ValGluTyrSerProValThrGluLysHisLeuThrAspGlyMetThrValArgGluLeu -

2221 TATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGA
 -----+-----+-----+-----+-----+-----+ 2280
 ATACGTCACGACGGTATTGGTACTCACTATTGTGACGCCGGTTGAATGAAGACTGTTGCT
 a: TyrAlaValLeuProEndProEndValIleThrLeuArgProThrTyrPheEndGlnArg -
 b: MetGlnCysCysHisAsnHisGluEndEndHisCysGlyGlnLeuThrSerAspAsnAsp -
 c: CysSerAlaAlaIleThrMetSerAspAsnThrAlaAlaAsnLeuLeuLeuThrThrIle -

P
V
U
I

2281 TCGGAGGACCGAAGGAGCTAACCGCTTTTTCACACAACATGGGGGATCATGTAAC TCGCC
-----+-----+-----+-----+-----+-----+-----+-----+ 2340
AGCCTCCTGGCTTCCTCGATTGGCGAAAAACGTGTTGTACCCCTAGTACATTGAGCGG

a: SerGluAspArgArgSerEndProLeuPheCysThrThrTrpGlyIleMetEndLeuAla -
b: ArgArgThrGluGlyAlaAsnArgPhePheAlaGlnHisGlyGlySerCysAsnSerPro -
c: GlyGlyProLysGluLeuThrAlaPheLeuHisAsnMetGlyAspHisValThrArgLeu -

2341 TTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGA
-----+-----+-----+-----+-----+-----+-----+-----+ 2400
AACTAGCAACCCCTTGGCCTCGACTTACTTCGGTATGGTTTGTCTGCTCGCACTGTGGTGCT

a: LeuIleValGlyAsnArgSerEndMetLysProTyrGlnThrThrSerValThrProArg -
b: EndSerLeuGlyThrGlyAlaGluEndSerHisThrLysArgArgAlaEndHisHisAsp -
c: AspArgTrpGluProGluLeuAsnGluAlaIleProAsnAspGluArgAspThrThrMet -

F
S
P
I

2401 TGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAAGTGGCGAACTACTTACTCTAG
-----+-----+-----+-----+-----+-----+-----+-----+ 2460
ACGGACATCGTTACCGTTGTTGCAACGCGTTTGATAATTGACCGCTTGATGAATGAGATC

a: CysLeuEndGlnTrpGlnGlnArgCysAlaAsnTyrEndLeuAlaAsnTyrLeuLeuEnd -
b: AlaCysSerAsnGlyAsnAsnValAlaGlnThrIleAsnTrpArgThrThrTyrSerSer -
c: ProValAlaMetAlaThrThrLeuArgLysLeuLeuThrGlyGluLeuLeuThrLeuAla -

2461 CTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCCTTCTGCG
-----+-----+-----+-----+-----+-----+-----+-----+ 2520
GAAGGGCCGTTGTTAATTATCTGACCTACCTCCGCCTATTTCAACGTCCTGGTGAAGACG

a: LeuProGlyAsnAsnEndEndThrGlyTrpArgArgIleLysLeuGlnAspHisPheCys -
b: PheProAlaThrIleAsnArgLeuAspGlyGlyGlyEndSerCysArgThrThrSerAla -
c: SerArgGlnGlnLeuIleAspTrpMetGluAlaAspLysValAlaGlyProLeuLeuArg -

B
g
l
I

2521 GCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT
-----+-----+-----+-----+-----+-----+-----+-----+ 2580
CGAGCCGGGAAGGCCGACCGACCAATAACGACTATTTAGACCTCGGCCACTCGCACCCA

a: AlaArgProPheArgLeuAlaGlyLeuLeuLeuIleAsnLeuGluProValSerValGly -
b: LeuGlyProSerGlyTrpLeuValTyrCysEndEndIleTrpSerArgEndAlaTrpVal -
c: SerAlaLeuProAlaGlyTrpPheIleAlaAspLysSerGlyAlaGlyGluArgGlySer -

B
s
a
I

P
f
l
l
l
0
8
I

2581 CTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCT
-----+-----+-----+-----+-----+ 2640
GAGCGCCATAGTAACGTCGTGACCCCGTCTACCATTCGGGAGGGCATAGCATCAATAGA

a: LeuAlaValSerLeuGlnHisTrpGlyGlnMetValSerProProValSerEndLeuSer -
b: SerArgTyrHisCysSerThrGlyAlaArgTrpEndAlaLeuProTyrArgSerTyrLeu -
c: ArgGlyIleIleAlaAlaLeuGlyProAspGlyLysProSerArgIleValValIleTyr -

E
a
m
l
l
0
5
I

2641 ACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTG
-----+-----+-----+-----+-----+ 2700
TGTGCTGCCCCCTCAGTCCGTTGATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCAC

a: ThrArgArgGlyValArgGlnLeuTrpMetAsnGluIleAspArgSerLeuArgEndVal -
b: HisAspGlyGluSerGlyAsnTyrGlyEndThrLysEndThrAspArgEndAspArgCys -
c: ThrThrGlySerGlnAlaThrMetAspGluArgAsnArgGlnIleAlaGluIleGlyAla -

2701 CCTCACTGATTAAGCATTGGTAAGTGTACACCAAGTTTACTCATATATACTTTAGATTG
-----+-----+-----+-----+-----+ 2760
GGAGTGACTAATTCGTAACCATTCAGACTCTGGTTCAAATGAGTATATATGAAATCTAAC

a: ProHisEndLeuSerIleGlyAsnCysGlnThrLysPheThrHisIleTyrPheArgLeu -
b: LeuThrAspEndAlaLeuValThrValArgProSerLeuLeuIleTyrThrLeuAspEnd -
c: SerLeuIleLysHisTrpEndLeuSerAspGlnValTyrSerTyrIleLeuEndIleAsp -

2761 ATTTAAACCTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTTTTGTATAATCTCA
-----+-----+-----+-----+-----+ 2820
TAAATTTTGAAGTAAAAATTAATTTTCTAGATCCACTTCTAGGAAAACTATTAGAGT

a: IleEndAsnPheIlePheAsnLeuLysGlySerArgEndArgSerPheLeuIleIleSer -
b: PheLysThrSerPheLeuIleEndLysAspLeuGlyGluAspProPheEndEndSerHis -
c: LeuLysLeuHisPheEndPheLysArgIleEndValLysIleLeuPheAspAsnLeuMet -

2821 TGACCAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGA
-----+-----+-----+-----+-----+ 2880
ACTGGTTTTAGGGAATTGCACTCAAAGCAAGGTGACTCGCAGTCTGGGGCATCTTTTCT

a: EndProLysSerLeuAsnValSerPheArgSerThrGluArgGlnThrProEndLysArg -
b: AspGlnAsnProLeuThrEndValPheValProLeuSerValArgProArgArgLysAsp -
c: ThrLysIleProEndArgGluPheSerPheHisEndAlaSerAspProValGluLysIle -

2881 TCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAA 2940
-----+-----+-----+-----+-----+-----+-----+-----+
AGTTTCCTAGAGAAGTCTAGGAAAAAAGACGCGCATTAGACGACGAACGTTTGTTTTTT

a: SerLysAspLeuLeuGluIleLeuPhePheCysAlaEndSerAlaAlaCysLysGlnLys -
b: GlnArgIlePheLeuArgSerPhePheSerAlaArgAsnLeuLeuLeuAlaAsnLysLys -
c: LysGlySerSerEndAspProPhePheLeuArgValIleCysCysLeuGlnThrLysLys -

2941 AACCACCGCTACCAGCGGTGGTTTGTGTCGGGATCAAGAGCTACCAACTCTTTTTCCGA 3000
-----+-----+-----+-----+-----+-----+-----+-----+
TTGGTGGCGATGGTCGCCACCAACAAACGGCCTAGTTCTCGATGGTTGAGAAAAAGGCT

a: AsnHisArgTyrGlnArgTrpPheValCysArgIleLysSerTyrGlnLeuPhePheArg -
b: ThrThrAlaThrSerGlyGlyLeuPheAlaGlySerArgAlaThrAsnSerPheSerGlu -
c: ProProLeuProAlaValValCysLeuProAspGlnGluLeuProThrLeuPheProLys -

3001 AGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCTTCTAGTGTAGCCGTAGT 3060
-----+-----+-----+-----+-----+-----+-----+-----+
TCCATTGACCGAAGTCGTCTCGCGTCTATGGTTTATGACAGGAAGATCACATCGGCATCA

a: ArgEndLeuAlaSerAlaGluArgArgTyrGlnIleLeuSerPheEndCysSerArgSer -
b: GlyAsnTrpLeuGlnGlnSerAlaAspThrLysTyrCysProSerSerValAlaValVal -
c: ValThrGlyPheSerArgAlaGlnIleProAsnThrValLeuLeuValEndProEndLeu -

3061 TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGT 3120
-----+-----+-----+-----+-----+-----+-----+-----+
ATCCGGTGGTGAAGTTCTTGAGACATCGTGGCGGATGTATGGAGCGAGACGATTAGGACA

a: EndAlaThrThrSerArgThrLeuEndHisArgLeuHisThrSerLeuCysEndSerCys -
b: ArgProProLeuGlnGluLeuCysSerThrAlaTyrIleProArgSerAlaAsnProVal -
c: GlyHisHisPheLysAsnSerValAlaProProThrTyrLeuAlaLeuLeuIleLeuLeu -

3121 TACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGAT 3180
-----+-----+-----+-----+-----+-----+-----+-----+
ATGGTCACCGACGACGGTCACCGCTATTTCAGCACAGAATGGCCCAACCTGAGTTCTGCTA

a: TyrGlnTrpLeuLeuProValAlaIleSerArgValLeuProGlyTrpThrGlnAspAsp -
b: ThrSerGlyCysCysGlnTrpArgEndValValSerTyrArgValGlyLeuLysThrIle -
c: ProValAlaAlaAlaSerGlyAspLysSerCysLeuThrGlyLeuAspSerArgArgEnd -

3181 AGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCT 3240
-----+-----+-----+-----+-----+-----+-----+-----+
TCAATGGCCTATTCCGCGTCGCCAGCCCGACTTGCCCCCAAGCACGTGTGTCGGGTGCA

a: SerTyrArgIleArgArgSerGlyArgAlaGluArgGlyValArgAlaHisSerProAla -
b: ValThrGlyEndGlyAlaAlaValGlyLeuAsnGlyGlyPheValHisThrAlaGlnLeu -
c: LeuProAspLysAlaGlnArgSerGlyEndThrGlyGlySerCysThrGlnProSerLeu -

3241 TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA 3300
-----+-----+-----+-----+-----+-----+-----+-----+
ACCTCGCTTGCTGGATGTGGCTTGACTCTATGGATGTGCGACTCGATACTCTTTTCGCGGT

a: TrpSerGluArgProThrProAsnEndAspThrTyrSerValSerTyrGluLysAlaPro -
b: GlyAlaAsnAspLeuHisArgThrGluIle ProThrAlaEndAlaMetArgLysArgHis -
c: GluArgThrThrTyrThrGluLeuArgTyrLeuGlnArgGluLeuEndGluSerAlaThr -

CGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAG
3301 -----+-----+-----+-----+-----+-----+-----+ 3360
GCGAAGGGCTTCCCTCTTTCCGCCTGTCCATAGGCCATTTCGCCGTCCCAGCCCTGTGTCCTC

a: ArgPheProLysGlyGluArgArgThrGlyIleArgEndAlaAlaGlySerGluGlnGlu -
b: AlaSerArgArgGluLysGlyGlyGlnValSerGlyLysArgGlnGlyArgAsnArgArg -
c: LeuProGluGlyArgLysAlaAspArgTyrProValSerGlyArgValGlyThrGlyGlu -

AGCGCACGAGGGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTCCGGCTTTC
3361 -----+-----+-----+-----+-----+-----+ 3420
TCGCGTGCTCCCTCGAAGGTCCCCCTTTGCGGACCATAGAAATATCAGGACAGCCCCAAG

a: SerAlaArgGlySerPheGlnGlyGluThrProGlyIlePheIleValLeuSerGlyPhe -
b: AlaHisGluGlyAlaSerArgGlyLysArgLeuValSerLeuEndSerCysArgValSer -
c: ArgThrArgGluLeuProGlyGlyAsnAlaTrpTyrLeuTyrSerProValGlyPheArg -

GCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGA
3421 -----+-----+-----+-----+-----+-----+ 3480
CGGTGGAGACTGAACTCGCAGCTAAAAACACTACGAGCAGTCCCCCGCCTCGGATACCT

a: AlaThrSerAspLeuSerValAspPheCysAspAlaArgGlnGlyGlyGlyAlaTyrGly -
b: ProProLeuThrEndAlaSerIlePheValMetLeuValArgGlyAlaGluProMetGlu -
c: HisLeuEndLeuGluArgArgPheLeuEndCysSerSerGlyGlyArgSerLeuTrpLys -

AAAACGCCAGCAACGCGGCCTTTTTACGGTTCTGGCCTTTTGCTGGCCTTTTGCTCACA
3481 -----+-----+-----+-----+-----+-----+ 3540
TTTTGCGGTCGTTGCGCCGAAAAATGCCAAGGACCGGAAACGACCGGAAACGAGTGT

a: LysThrProAlaThrArgProPheTyrGlySerTrpProPheAlaGlyLeuLeuLeuThr -
b: LysArgGlnGlnArgGlyLeuPheThrValProGlyLeuLeuLeuAlaPheCysSerHis -
c: AsnAlaSerAsnAlaAlaPheLeuArgPheLeuAlaPheCysTrpProPheAlaHisMet -

TGTTCTTTCTTCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAG
3541 -----+-----+-----+-----+-----+-----+ 3600
ACAAGAAAGGACGCAATAGGGGACTAAGACACCTATTGGCATAATGGCGGAAACTCACTC

a: CysSerPheLeuArgTyrProLeuIleLeuTrpIleThrValLeuProProLeuSerGlu -
b: ValLeuSerCysValIleProEndPheCysGlyEndProTyrTyrArgLeuEndValSer -
c: PhePheProAlaLeuSerProAspSerValAspAsnArgIleThrAlaPheGluEndAla -

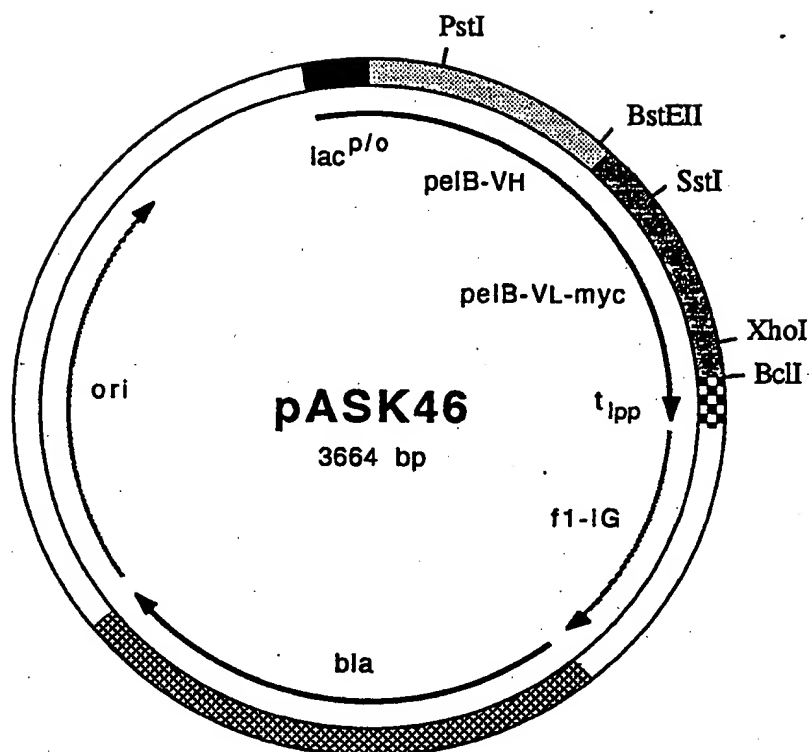
CTGATACCGCTCGCCGACGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGG
3601 -----+-----+-----+-----+-----+-----+ 3660
GACTATGGCGAGCGGCGTCCGCTTGCTGGCTCGCGTCGCTCAGTCACTCGCTCCTTCGCC

a: LeuIleProLeuAlaAlaAlaGluArgProSerAlaAlaSerGlnEndAlaArgLysArg -
b: EndTyrArgSerProGlnProAsnAspArgAlaGlnArgValSerGluArgGlySerGly -
c: AspThrAlaArgArgSerArgThrThrGluArgSerGluSerValSerGluGluAlaGlu -

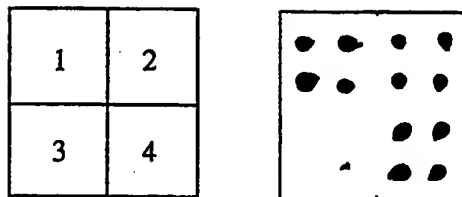
AAGA
3661 ----- 3664
TTCT

a: Lys -
b: Arg -
c: -

Figur 2



Figur 3



Figur 4

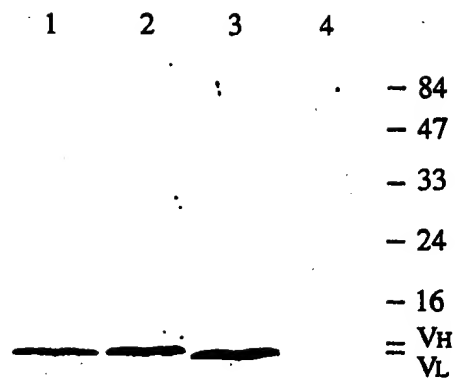
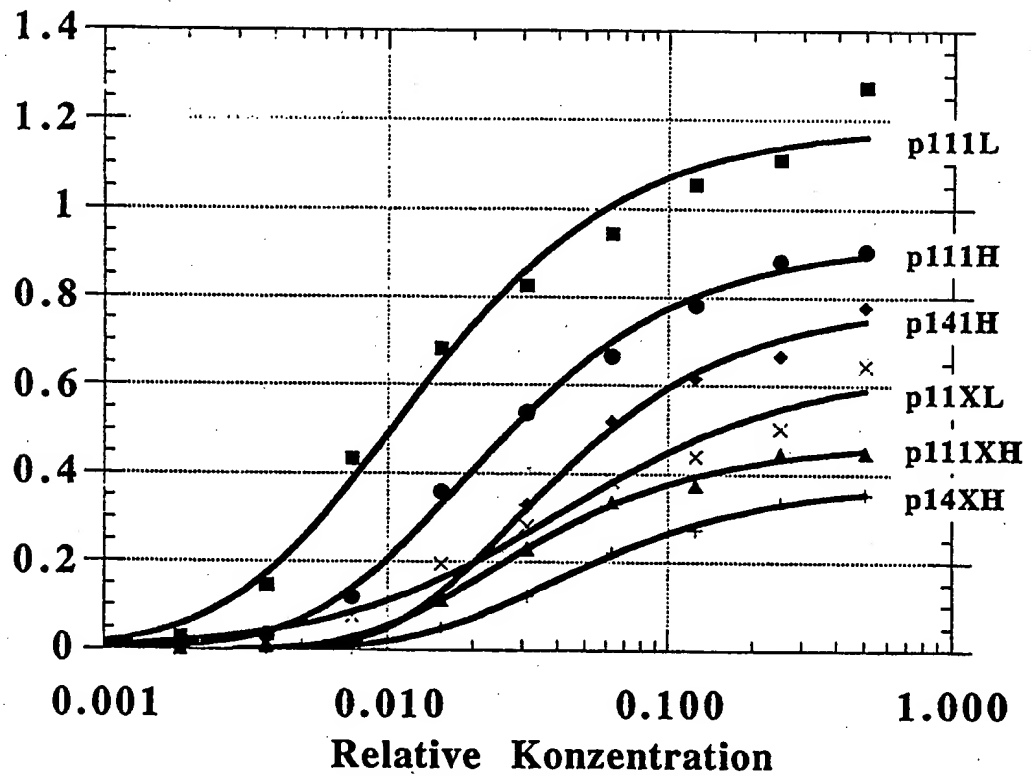


Figure 5

A405-A450

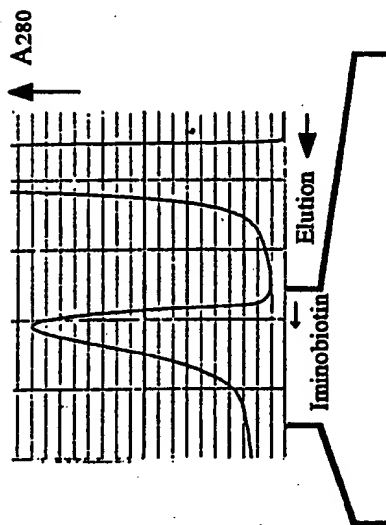


Nummer:
Int. Cl.⁵:
Offenlegungstag:

DE 42 37 113 A1
C 07 K 7/06
5. Mai 1994

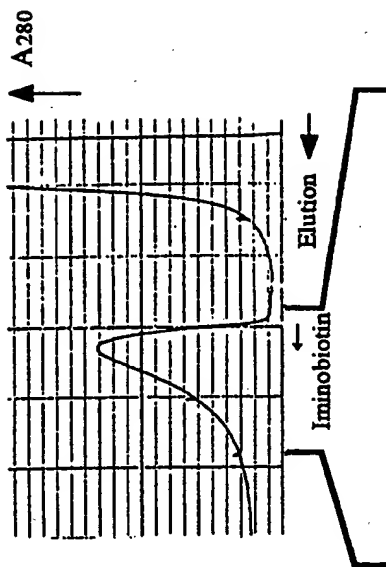
Figur 6A

p111H



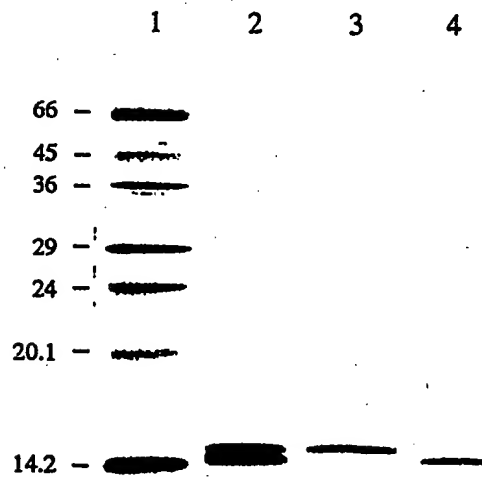
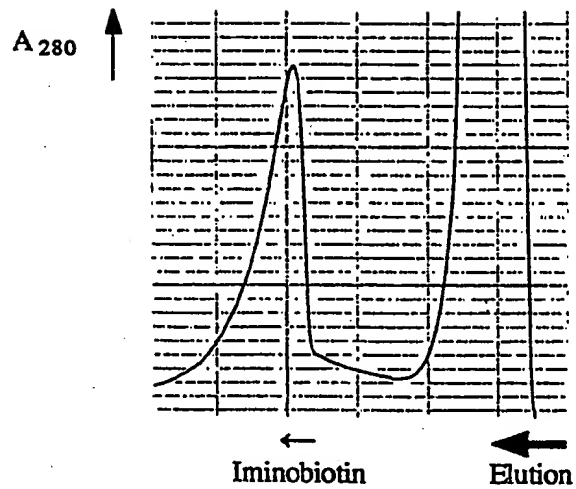
$\frac{V_L}{V_H}$

p111L

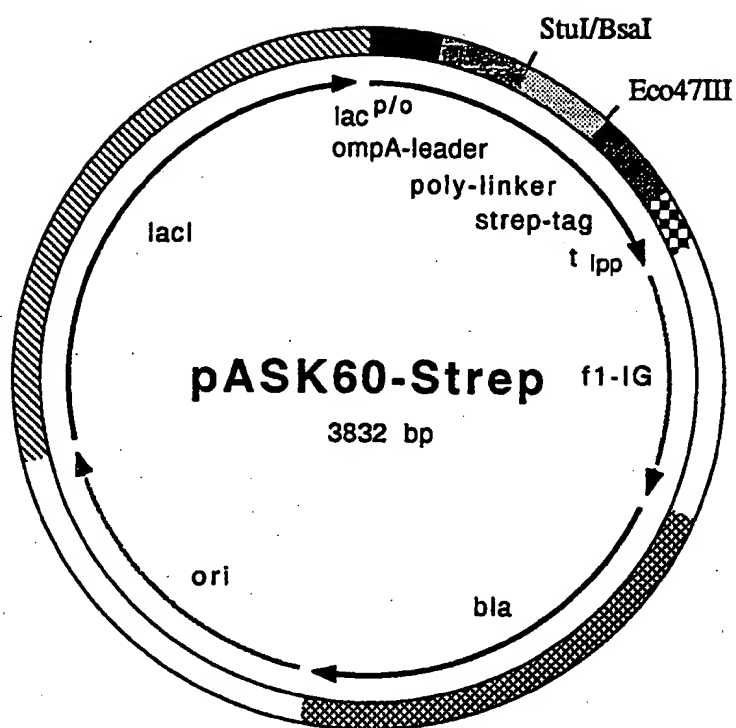


$\frac{V_L}{V_H}$

Figur 6B



Figur 7



Figur 8

DNS-Sequenz des Plasmids pASK60-Strep

Gezeigt sind Strang (mit kodierte Aminosäuresequenzen in allen drei Leserahmen) und Gegenstrang mit singulären Restriktionsschnittstellen.

P
f
l
M
I

```

ACCCGACACCATCGAATGGCGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGA
1  -----+-----+-----+-----+-----+-----+-----+-----+ 60
TGGGCTGTGGTAGCTTACCGCGTTTGGAAAGCGCCATACCGTACTATCGCGGGCCTTCT

a:   ThrArgHisHisArgMetAlaGlnAsnLeuSerArgTyrGlyMetIleAlaProGlyArg -
b:   ProAspThrIleGluTrpArgLysThrPheArgGlyMetAlaEndEndArgProGluGlu -
c:   ProThrProSerAsnGlyAlaLysProPheAlaValTrpHisAspSerAlaArgLysArg -

GAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGC
61 -----+-----+-----+-----+-----+-----+-----+-----+ 120
CTCAGTTAAGTCCACCACTTACACTTTGGTCATTGCAATATGCTACAGCGTCTCATACG

a:   GluSerIleGlnGlyGlyGluCysGluThrSerAsnValIleArgCysArgArgValCys -
b:   SerGlnPheArgValValAsnValLysProValThrLeuTyrAspValAlaGluTyrAla -
c:   ValAsnSerGlyTrpEndMetEndAsnGlnEndArgTyrThrMetSerGlnSerMetPro -

D
r
d
I
I
I

CGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAA
121 -----+-----+-----+-----+-----+-----+-----+-----+ 180
GCCACAGAGAATAGTCTGGCAAAGGGCGCACCCTTGGTCCGGTCCGTCAAAGACGCTT

a:   ArgCysLeuLeuSerAspArgPheProArgGlyGluProGlyGlnProArgPheCysGlu -
b:   GlyValSerTyrGlnThrValSerArgValValAsnGlnAlaSerHisValSerAlaLys -
c:   ValSerLeuIleArgProPheProAlaTrpEndThrArgProAlaThrPheLeuArgLys -

AACGCGGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGC
181 -----+-----+-----+-----+-----+-----+-----+-----+ 240
TTGCGCCCTTTTTCACCTTCGCCGCTACCGCCTCGACTTAATGTAAGGTTGGCGCACCG

a:   AsnAlaGlyLysSerGlySerGlyAspGlyGlyAlaGluLeuHisSerGlnProArgGly -
b:   ThrArgGluLysValGluAlaAlaMetAlaGluLeuAsnTyrIleProAsnArgValAla -
c:   ArgGlyLysLysTrpLysArgArgTrpArgSerEndIleThrPheProThrAlaTrpHis -

ACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCT
241 -----+-----+-----+-----+-----+-----+-----+-----+ 300
TGTTGTTGACCGCCCGTTTGTCTAGCAACGACTAACCGCAACGGTGGAGGTCAGACCGGGA

a:   ThrThrThrGlyGlyGlnThrValValAlaAspTrpArgCysHisLeuGlnSerGlyPro -
b:   GlnGlnLeuAlaGlyLysGlnSerLeuLeuIleGlyValAlaThrSerSerLeuAlaLeu -
c:   AsnAsnTrpArgAlaAsnSerArgCysEndLeuAlaLeuProProProValTrpProCys -

```

GCACGCGCCGTCGCAAATTTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAG
 301 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 360
 CGTGCGCGGCAGCGTTTAAACAGCGCCGCTAATTTAGAGCGCGGCTAGTTGACCCACGGTC

a: AlaArgAlaValAlaAsnCysArgGlyAspEndIleSerArgArgSerThrGlyCysGln -
 b: HisAlaProSerGlnIleValAlaAlaIleLysSerArgAlaAspGlnLeuGlyAlaSer -
 c: ThrArgArgArgLysLeuSerArgArgLeuAsnLeuAlaProIleAsnTrpValProAla -

CGTGCTGGTGTCTGATGGTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGGCGGTGCACAA
 361 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 420
 GCACCACCACAGCTACCATCTTGCTTCGCCCGAGCTTCGGACATTTCCGCCCCACGTGTT

a: ArgGlyGlyValAspGlyArgThrLysArgArgArgSerLeuEndSerGlyGlyAlaGln -
 b: ValValValSerMetValGluArgSerGlyValGluAlaCysLysAlaAlaValHisAsn -
 c: TrpTrpCysArgTrpEndAsnGluAlaAlaSerLysProValLysArgArgCysThrIle -

A
 f
 lM B
 Il C
 lu l
 II I
 /

TCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAGGATGC
 421 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 480
 AGAAGAGCGCGTTGCGCAGTCACCCGACTAGTAATTGATAGGCGACCTACTGGTCCTACG

a: SerSerArgAlaThrArgGlnTrpAlaAspHisEndLeuSerAlaGlyEndProGlyCys -
 b: LeuLeuAlaGlnArgValSerGlyLeuIleIleAsnTyrProLeuAspAspGlnAspAla -
 c: PheSerArgAsnAlaSerValGlyEndSerLeuThrIleArgTrpMetThrArgMetPro -

CATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCA
 481 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 540
 GTAACGACACCTTCGACGGACGTGATTACAAGGCCGCAATAAAGAAGTACAGAGACTGGT

a: HisCysCysGlySerCysLeuHisEndCysSerGlyValIleSerEndCysLeuEndPro -
 b: IleAlaValGluAlaAlaCysThrAsnValProAlaLeuPheLeuAspValSerAspGln -
 c: LeuLeuTrpLysLeuProAlaLeuMetPheArgArgTyrPheLeuMetSerLeuThrArg -

GACACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGAGCA
 541 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 600
 CTGTGGGTAGTTGTCATAATAAAGAGGGTACTTCTGCCATGCGCTGACCCGCACCTCGT

a: AspThrHisGlnGlnTyrTyrPheLeuProEndArgArgTyrAlaThrGlyArgGlyAla -
 b: ThrProIleAsnSerIleIlePheSerHisGluAspGlyThrArgLeuGlyValGluHis -
 c: HisProSerThrValLeuPheSerProMetLysThrValArgAspTrpAlaTrpSerIle -

A
 p
 a
 I

TCTGGTTCGCATTGGGTCATCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTC
 601 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 660
 AGACCAGCGTAACCCAGTAGTCGTTTAGCGCGACAATCGCCCGGGTAATCAAGACAGAG

a: SerGlyArgIleGlySerSerAlaAsnArgAlaValSerGlyProIleLysPheCysLeu - -
 b: LeuValAlaLeuGlyHisGlnGlnIleAlaLeuLeuAlaGlyProLeuSerSerValSer - -
 c: TrpSerHisTrpValIleSerLysSerArgCysEndArgAlaHisEndValLeuSerArg - -

GGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGAT
661 -----+-----+-----+-----+-----+-----+-----+ 720
CCGCGCAGACGCAGACCGACCGACCGTATTTATAGAGTGAGCGTTAGTTTAAGTCGGCTA

a: GlyAlaSerAlaSerGlyTrpLeuAlaEndIleSerHisSerGlnSerAsnSerAlaAsp -
b: AlaArgLeuArgLeuAlaGlyTrpHisLysTyrLeuThrArgAsnGlnIleGlnProIle -
c: ArgValCysValTrpLeuAlaGlyIleAsnIleSerLeuAlaIleLysPheSerArgEnd -

AGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCT
721 -----+-----+-----+-----+-----+-----+-----+ 780
TCGCCTTGCCCTTCCGCTGACCTCACGGTACAGGCCAAAAGTTGTTTGGTACGTTTACGA

a: SerGlyThrGlyArgArgLeuGluCysHisValArgPheSerThrAsnHisAlaAsnAla -
b: AlaGluArgGluGlyAspTrpSerAlaMetSerGlyPheGlnGlnThrMetGlnMetLeu -
c: ArgAsnGlyLysAlaThrGlyValProCysProValPheAsnLysProCysLysCysEnd -

GAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGC
781 -----+-----+-----+-----+-----+-----+-----+ 840
CTTACTCCCGTAGCAAGGGTGACGCTACGACCAACGGTTGCTAGTCTACCGCGACCCGCG

a: GluEndGlyHisArgSerHisCysAspAlaGlyCysGlnArgSerAspGlyAlaGlyArg -
b: AsnGluGlyIleValProThrAlaMetLeuValAlaAsnAspGlnMetAlaLeuGlyAla -
c: MetArgAlaSerPheProLeuArgCysTrpLeuProThrIleArgTrpArgTrpAlaGln -

B
S
S
H
I
I
E
C
O
R
V

AATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTCGGTAGTGGGATA
841 -----+-----+-----+-----+-----+-----+-----+ 900
TTACGCGCGGTAATGGCTCAGGCCCCGACGCGCAACCACGCCTATAGAGCCATCACCCCTAT

a: AsnAlaArgHisTyrArgValArgAlaAlaArgTrpCysGlyTyrLeuGlySerGlyIle -
b: MetArgAlaIleThrGluSerGlyLeuArgValGlyAlaAspIleSerValValGlyTyr -
c: CysAlaProLeuProSerProGlyCysAlaLeuValArgIleSerArgEndTrpAspThr -

H
P
a
I

CGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTT
901 -----+-----+-----+-----+-----+-----+-----+ 960
GCTGCTATGGCTTCTGTGAGTACAATATAGGGCGGCAATTGGTGGTAGTTTGTCTCTAAA

a: ArgArgTyrArgArgGlnLeuMetLeuTyrProAlaValAsnHisHisGlnThrGlyPhe -
b: AspAspThrGluAspSerSerCysTyrIleProProLeuThrThrIleLysGlnAspPhe -
c: ThrIleProLysThrAlaHisValIleSerArgArgEndProProSerAsnArgIlePhe -

TCGCCTGCTGGGGCAAACCGAGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGT
961 -----+-----+-----+-----+-----+-----+-----+ 1020
AGCGGACGACCCCGTTTGGTCGCACCTGGCGAACGACGTTGAGAGAGTCCCGGTCCGCCA

a: SerProAlaGlyAlaAsnGlnArgGlyProLeuAlaAlaThrLeuSerGlyProGlyGly -
b: ArgLeuLeuGlyGlnThrSerValAspArgLeuLeuGlnLeuSerGlnGlyGlnAlaVal -
c: AlaCysTrpGlyLysProAlaTrpThrAlaCysCysAsnSerLeuArgAlaArgArgEnd -

E
S
P
3
I

N
a
r
I

1021 GAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAA 1080
-----+-----+-----+-----+-----+-----+
CTTCCCGTTAGTCGACAACGGGCAGAGTGACCACTTTTCTTTTGGTGGGACCGCGGGTT

a: GluGlyGlnSerAlaValAlaArgLeuThrGlyGluLysLysAsnHisProGlyAlaGln -
b: LysGlyAsnGlnLeuLeuProValSerLeuValLysArgLysThrThrLeuAlaProAsn -
c: ArgAlaIleSerCysCysProSerHisTrpEndLysGluLysProProTrpArgProIle -

1081 TACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGT 1140
-----+-----+-----+-----+-----+-----+
ATGCGTTTGGCGGAGAGGGGCGCGCAACCGGCTAAGTAATTACGTCGACCGTGCTGTCCA

a: TyrAlaAsnArgLeuSerProArgValGlyArgPheIleAsnAlaAlaGlyThrThrGly -
b: ThrGlnThrAlaSerProArgAlaLeuAlaAspSerLeuMetGlnLeuAlaArgGlnVal -
c: ArgLysProProLeuProAlaArgTrpProIleHisEndCysSerTrpHisAspArgPhe -

1141 TTCCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATT 1200
-----+-----+-----+-----+-----+-----+
AAGGGCTGACCTTTTCGCCCGTCACTCGCGTTGCGTTAATTACACTCAATCGAGTGAGTAA

a: PheProThrGlyLysArgAlaValSerAlaThrGlnLeuMetEndValSerSerLeuIle -
b: SerArgLeuGluSerGlyGlnEndAlaGlnArgAsnEndCysGluLeuAlaHisSerLeu -
c: ProAspTrpLysAlaGlySerGluArgAsnAlaIleAsnValSerEndLeuThrHisEnd -

1201 AGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCG 1260
-----+-----+-----+-----+-----+-----+
TCCGTGGGGTCCGAAATGTGAAATACGAAGGCCGAGCATACAACACACCTTAACACTCGC

a: ArgHisProArgLeuTyrThrLeuCysPheArgLeuValCysCysValGluLeuEndAla -
b: GlyThrProGlyPheThrLeuTyrAlaSerGlySerTyrValValTrpAsnCysGluArg -
c: AlaProGlnAlaLeuHisPheMetLeuProAlaArgMetLeuCysGlyIleValSerGly -

X
b
a
I

1261 GATAACAATTTACACAGGAAACAGCTATGACCATGATTACGAATTTCTAGATAACGAGG 1320
-----+-----+-----+-----+-----+-----+
CTATTGTTAAAGTGTGTCCTTTGTGCGATACTGGTACTAATGCTTAAAGATCTATTGCTCC

a: AspAsnAsnPheThrGlnGluThrAlaMetThrMetIleThrAsnPheEndIleThrArg -
b: IleThrIleSerHisArgLysGlnLeuEndProEndLeuArgIleSerArgEndArgGly -
c: EndGlnPheHisThrGlyAsnSerTyrAspHisAspTyrGluPheLeuAspAsnGluGly -

N
r
u
I

1321 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1380
GCAAAAATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCTACCGT
CGTTTTTTTACTTTTTCTGTGTCGATAGCGCTAACGTCACCGTGACCGACCAAAGCGATGGCA

a: AlaLysAsnGluLysAspSerTyrArgAspCysSerGlyThrGlyTrpPheArgTyrArg -
b: GlnLysMetLysLysThrAlaIleAlaIleAlaValAlaLeuAlaGlyPheAlaThrVal -
c: LysLysEndLysArgGlnLeuSerArgLeuGlnTrpHisTrpLeuValSerLeuProEnd -

1381 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1440
AGCGCAGGCTGAGACCAGAAATTCGAGCTCGGTACCCGGGGATCCCTCGAGGTTCGACCTG
TCGCGTCCGGACTCTGGTCTTAAGCTCGAGCCATGGGCCCTAGGGAGCTCCAGCTGGAC

a: SerAlaGlyLeuArgProGluPheGluLeuGlyThrArgGlySerLeuGluValAspLeu -
b: AlaGlnAlaEndAspGlnAsnSerSerSerValProGlyAspProSerArgSerThrCys -
c: ArgArgProGluThrArgIleArgAlaArgTyrProGlyIleProArgGlyArgProAla -

S E
s c
e o
8 B 4
P3 s 7
s8 p I
t7 M I
II I I
/

1441 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1500
CAGGCAGCGCTTGGCGTCACCCGCGAGTTCGGTGGTTAATAAGCTTGACCTGTGAAGTGAA
GTCCGTCGCGAACCAGCGAGTGGGCGTCAAGCCACCAATTATTGGAAGTGGACACTTCACIT

a: GlnAlaAlaLeuGlyValThrArgSerSerValValAsnLysLeuAspLeuEndSerGlu -
b: ArgGlnArgLeuAlaSerProAlaValArgTrpLeuIleSerLeuThrCysGluValLys -
c: GlySerAlaTrpArgHisProGlnPheGlyGlyEndEndAlaEndProValLysEndLys -

1501 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1560
AAATGGCGCACATTGTGCGACATTTTTTTTGTCTGCCGTTTACCGCTACTGCGTCACGGA
TTTACCGCGTGTAAACAGCTGTAAAAAAACAGACGGCAAATGGCGATGACGCGAGTGCCT

a: LysTrpArgThrLeuCysAspIlePhePheValCysArgLeuProLeuLeuArgHisGly -
b: AsnGlyAlaHisCysAlaThrPhePheLeuSerAlaValTyrArgTyrCysValThrAsp -
c: MetAlaHisIleValArgHisPhePheCysLeuProPheThrAlaThrAlaSerArgIle -

1561 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1620
TCTCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGTGTGGTGGTTACGCGCAGCGT
AGAGGTGCGCGGACATCGCCGCGTAATTCGCGCGCCACACCACCAATGCGCGTTCGCA

a: SerProArgAlaLeuEndArgArgIleLysArgGlyGlyCysGlyGlyTyrAlaGlnArg -
b: LeuHisAlaProCysSerGlyAlaLeuSerAlaAlaGlyValValValThrArgSerVal -
c: SerThrArgProValAlaAlaHisEndAlaArgArgValTrpTrpLeuArgAlaAlaEnd -

1621 GACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTTCGCTTTCTTCCCTTCCTTTCT
 -----+-----+-----+-----+-----+-----+-----+ 1680
 CTGGCGATGTGAACGGTCGCGGGATCGCGGGCGAGGAAAGCGAAAGAAGGGAAGGAAAGA

 a: AspArgTyrThrCysGlnArgProSerAlaArgSerPheArgPheLeuProPheLeuSer -
 b: ThrAlaThrLeuAlaSerAlaLeuAlaProAlaProPheAlaPhePheProSerPheLeu -
 c: ProLeuHisLeuProAlaProEndArgProLeuLeuSerLeuSerSerLeuProPheSer -

 N
 a
 e
 I
 1681 CGCCACGTTTCGCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCG
 -----+-----+-----+-----+-----+-----+ 1740
 GCGGTGCAAGCGGCCGAAAGGGGCAGTTTCGAGATTTAGCCCCCGAGGGAAATCCCAAGGC

 a: ArgHisValArgArgLeuSerProSerSerSerLysSerGlyAlaProPheArgValPro -
 b: AlaThrPheAlaGlyPheProArgGlnAlaLeuAsnArgGlyLeuProLeuGlyPheArg -
 c: ProArgSerProAlaPheProValLysLeuEndIleGlyGlySerLeuEndGlySerAsp -

 B
 S
 a
 A
 I
 1741 ATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTACGCTAG
 -----+-----+-----+-----+-----+-----+ 1800
 TAAATCACGAAATGCCGTGGAGCTGGGGTTTTTTTGAACATAATCCCACTACCAAGTGCATC

 a: IleEndCysPheThrAlaProArgProGlnLysThrEndLeuGlyEndTrpPheThrEnd -
 b: PheSerAlaLeuArgHisLeuAspProLysLysLeuAspEndGlyAspGlySerArgSer -
 c: LeuValLeuTyrGlyThrSerThrProLysAsnLeuIleArgValMetValHisValVal -

 TGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAA
 1801 -----+-----+-----+-----+-----+-----+ 1860
 ACCCGGTAGCGGGACTATCTGCCAAAAAGCGGAAACTGCAACCTCAGGTGCAAGAAATT

 a: TrpAlaIleAlaLeuIleAspGlyPheSerProPheAspValGlyValHisValLeuEnd -
 b: GlyProSerProEndEndThrValPheArgProLeuThrLeuGluSerThrPhePheAsn -
 c: GlyHisArgProAspArgArgPhePheAlaLeuEndArgTrpSerProArgSerLeuIle -

 TAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGA
 1861 -----+-----+-----+-----+-----+-----+ 1920
 ATCACCTGAGAACAAGGTTTGACCTTGTTGTGAGTTGGGATAGAGCCAGATAAGAAAAC

 a: EndTrpThrLeuValProAsnTrpAsnAsnThrGlnProTyrLeuGlyLeuPhePheEnd -
 b: SerGlyLeuLeuPheGlnThrGlyThrThrLeuAsnProIleSerValTyrSerPheAsp -
 c: ValAspSerCysSerLysLeuGluGlnHisSerThrLeuSerArgSerIleLeuLeuIle -

 TTTATAAGGGATTTTGCCGATTTTCGGCTATTTGGTTAAAAATGAGCTGATTTAACAAAA
 1921 -----+-----+-----+-----+-----+-----+ 1980
 AAATATTTCCCTAAACGGCTAAAGCCGGATAACCAATTTTTTACTCGACTAAATTGTTTT

 a: PheIleArgAspPheAlaAspPheGlyLeuLeuValLysLysEndAlaAspLeuThrLys -
 b: LeuEndGlyIleLeuProIleSerAlaTyrTrpLeuLysAsnGluLeuIleEndGlnLys -
 c: TyrLysGlyPheCysArgPheArgProIleGlyEndLysMetSerEndPheAsnLysAsn -

1981 ATTTAACGCGAATTTTAAACAAAATATTAAACGTTTACAATTTTCAGGTGGCACTTTTCGGGG
-----+-----+-----+-----+-----+ 2040
TAAATTGCGCTTAAATTTGTTTTATAATTGCAAATGTTAAAGTCCACCGTGAAAAGCCCC

a: IleEndArgGluPheEndGlnAsnIleAsnValTyrAsnPheArgTrpHisPheSerGly -
b: PheAsnAlaAsnPheAsnLysIleLeuThrPheThrIleSerGlyGlyThrPheArgGly -
c: LeuThrArgIleLeuThrLysTyrEndArgLeuGlnPheGlnValAlaLeuPheGlyGlu -

2041 AAATGTGCGCGGAACCCCTATTTGTTTTATTTTCTAAATACATTCAAATATGTATCCGCT
-----+-----+-----+-----+-----+ 2100
TTTACACGCGCCTTGGGGATAAACAAATAAAAAGATTTATGTAAGTTTATACATAGGCCGA

a: LysCysAlaArgAsnProTyrLeuPheIlePheLeuAsnThrPheLysTyrValSerAla -
b: AsnValArgGlyThrProIleCysLeuPhePheEndIleHisSerAsnMetTyrProLeu -
c: MetCysAlaGluProLeuPheValTyrPheSerLysTyrIleGlnIleCysIleArgSer -

2101 CATGAGACAATAACCCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTAT
-----+-----+-----+-----+-----+ 2160
GTACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCCTTCTCATACTCATA

a: HisGluThrIleThrLeuIleAsnAlaSerIleIleLeuLysLysGluGluTyrGluTyr -
b: MetArgGlnEndProEndEndMetLeuGlnEndTyrEndLysArgLysSerMetSerIle -
c: EndAspAsnAsnProAspLysCysPheAsnAsnIleGluLysGlyArgValEndValPhe -

2161 TCAACATTTCCGTGTCGCCCTTATTCCCTTTTTCGCGGCATTTTGCCTTCCTGTTTTCG
-----+-----+-----+-----+-----+ 2220
AGTTGTAAAGGCACAGCGGAATAAGGGAAAAACGCCGTAAACGGAAGGACAAAAACG

a: SerThrPheProCysArgProTyrSerLeuPheCysGlyIleLeuProSerCysPheCys -
b: GlnHisPheArgValAlaLeuIleProPhePheAlaAlaPheCysLeuProValPheAla -
c: AsnIleSerValSerProLeuPheProPheLeuArgHisPheAlaPheLeuPheLeuLeu -

2221 TCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGG
-----+-----+-----+-----+-----+ 2280
AGTGGGTCTTTGCGACCACCTTTCATTTTCTACGACTTCTAGTCAACCCACGTGCTCACCC

a: SerProArgAsnAlaGlyGluSerLysArgCysEndArgSerValGlyCysThrSerGly -
b: HisProGluThrLeuValLysValLysAspAlaGluAspGlnLeuGlyAlaArgValGly -
c: ThrGlnLysArgTrpEndLysEndLysMetLeuLysIleSerTrpValHisGluTrpVal -

2281 TTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACG
-----+-----+-----+-----+-----+ 2340
AATGTAGCTTGACCTAGAGTTGTCGCCATTCTAGGAATCTCAAAGCGGGGCTTCTTCG

a: LeuHisArgThrGlySerGlnGlnArgEndAspProEndGluPheSerProArgArgThr -
b: TyrIleGluLeuAspLeuAsnSerGlyLysIleLeuGluSerPheArgProGluGluArg -
c: ThrSerAsnTrpIleSerThrAlaValArgSerLeuArgValPheAlaProLysAsnVal -

2341 TTTTCCATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGA
-----+-----+-----+-----+-----+ 2400
AAAAGGTTACTACTCGTGAAAAATTTCAAGACGATACACCGCGCCATAATAGGGCATAACT

a: PheSerAsnAspGluHisPheEndSerSerAlaMetTrpArgGlyIleIleProTyrEnd -
b: PheProMetMetSerThrPheLysValLeuLeuCysGlyAlaValLeuSerArgIleAsp -
c: PheGlnEndEndAlaLeuLeuLysPh CysTyrValAlaArgTyrTyrProValLeuThr -

X
m
n
I

S
C
a
I

CGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTA
 2401 -----+-----+-----+-----+ 2460
 GCGGCCCGTTCTCGTTGAGCCAGCGCGTATGTGATAAGAGTCTTACTGAACCAACTCAT

a: ArgArgAlaArgAlaThrArgSerProHisThrLeuPheSerGluEndLeuGlyEndVal -
 b: AlaGlyGlnGluGlnLeuGlyArgArgIleHisTyrSerGlnAsnAspLeuValGluTyr -
 c: ProGlyLysSerAsnSerValAlaAlaTyrThrIleLeuArgMetThrTrpLeuSerThr -

CTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGC
 2461 -----+-----+-----+-----+ 2520
 GAGTGGTCAGTGTCTTTTCGTAGAATGCCTACCGTACTGTCTATTCTTAAATACGTCACG

a: LeuThrSerHisArgLysAlaSerTyrGlyTrpHisAspSerLysArgIleMetGlnCys -
 b: SerProValThrGluLysHisLeuThrAspGlyMetThrValArgGluLeuCysSerAla -
 c: HisGlnSerGlnLysSerIleLeuArgMetAlaEndGlnEndGluAsnTyrAlaValLeu -

P
V
u
I

TGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACC
 2521 -----+-----+-----+-----+ 2580
 ACGGTATTGGTACTCACTATTGTGACGCCGGTTGAATGAAGACTGTTGCTAGCCTCCTGG

a: CysHisAsnHisGluEndEndHisCysGlyGlnLeuThrSerAspAsnAspArgArgThr -
 b: AlaIleThrMetSerAspAsnThrAlaAlaAsnLeuLeuLeuThrThrIleGlyGlyPro -
 c: ProEndProEndValIleThrLeuArgProThrTyrPheEndGlnArgSerGluAspArg -

GAAGGAGCTAACCGCTTTTTCGACAAACATGGGGGATCATGTAACCTCGCCTTGATCGTTG
 2581 -----+-----+-----+-----+ 2640
 CTTCTCGATTGGCGAAAAACGTGTTGTACCCCTAGTACATTGAGCGGAAGTACGCAAC

a: GluGlyAlaAsnArgPhePheAlaGlnHisGlyGlySerCysAsnSerProEndSerLeu -
 b: LysGluLeuThrAlaPheLeuHisAsnMetGlyAspHisValThrArgLeuAspArgTrp -
 c: ArgSerEndProLeuPheCysThrThrTrpGlyIleMetEndLeuAlaLeuIleValGly -

GGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGC
 2641 -----+-----+-----+-----+ 2700
 CCTTGGCCTCGACTTACTTCGGTATGTTTGCTGCTCGCACTGTGGTGCTACGGACATCG

a: GlyThrGlyAlaGluEndSerHisThrLysArgArgAlaEndHisHisAspAlaCysSer -
 b: GluProGluLeuAsnGluAlaIleProAsnAspGluArgAspThrThrMetProValAla -
 c: AsnArgSerEndMetLysProTyrGlnThrThrSerValThrProArgCysLeuEndGln -

F
S
P
I

AATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCA
 2701 -----+-----+-----+-----+ 2760
 TTACCGTTGTTGCAACGCGTTTGATAATTGACCGCTTGATGAATGAGATCGAAGGGCCGT

a: AsnGlyAsnAsnValAlaGlnThrIleAsnTrpArgThrThrTyrS rS rPheProAla -
 b: MetAlaThrThrLeuArgLysLeuLeuThrGlyGluLeuLeuThrLeuAlaSerArgGln -
 c: TrpGlnGlnArgCysAlaAsnTyrEndLeuAlaAsnTyrLeuLeuEndLeuProGlyAsn -

2761 ACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2820
 TGTTAATTATCTGACCTACCTCCGCCTATTTCAACGTCCTGGTGAAGACGCGAGCCGGGA

a: ThrIleAsnArgLeuAspGlyGlyGlyEndSerCysArgThrThrSerAlaLeuGlyPro -
 b: GlnLeuIleAspTrpMetGluAlaAspLysValAlaGlyProLeuLeuArgSerAlaLeu -
 c: AsnEndEndThrGlyTrpArgArgIleLysLeuGlnAspHisPheCysAlaArgProPhe -

B
 g
 l
 I
 2821 TCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGCTCTCGCGGTAT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2880
 AGGCCGACCGACCAAATAACGACTATTTAGACCTCGGCCACTCGCACCGAGAGCGCCATA

a: SerGlyTrpLeuValTyrCysEndEndIleTrpSerArgEndAlaTrpLeuSerArgTyr -
 b: ProAlaGlyTrpPheIleAlaAspLysSerGlyAlaGlyGluArgGlySerArgGlyIle -
 c: ArgLeuAlaGlyLeuLeuLeuIleAsnLeuGluProValSerValAlaLeuAlaValSer -

P	E
f	a
l	m
l	l
l	l
0	0
8	5
I	I

2881 CATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2940
 GTAACGTCGTGACCCCGGTCTACCATTCCGGGAGGCGATAGCATCAATAGATGTGCTGCCCC

a: HisCysSerThrGlyAlaArgTrpEndAlaLeuProTyrArgSerTyrLeuHisAspGly -
 b: IleAlaAlaLeuGlyProAspGlyLysProSerArgIleValValIleTyrThrThrGly -
 c: LeuGlnHisTrpGlyGlnMetValSerProProValSerEndLeuSerThrArgArgGly -

2941 GAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGAT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3000
 CTCAGTCCGTTGATACCTACTTGTCTTTATCTGTCTAGCGACTCTATCCACGGAGTGACTA

a: GluSerGlyAsnTyrGlyEndThrLysEndThrAspArgEndAspArgCysLeuThrAsp -
 b: SerGlnAlaThrMetAspGluArgAsnArgGlnIleAlaGluIleGlyAlaSerLeuIle -
 c: ValArgGlnLeuTrpMetAsnGluIleAspArgSerLeuArgEndValProHisEndLeu -

3001 TAAGCATTGGTAACTGTGACACCAAGTTTACTCATATATACTTTAGATTGATTTAAACT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3060
 ATTCGTAACCATGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTAAATTTTGA

a: EndAlaLeuValThrValArgProSerLeuLeuIleTyrThrLeuAspEndPheLysThr -
 b: LysHisTrpEndLeuSerAspGlnValTyrSerTyrIleLeuEndIleAspLeuLysLeu -
 c: SerIleGlyAsnCysGlnThrLysPheThrHisIleTyrPheArgLeuIleEndAsnPhe -

3061 TCATTTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAAT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3120
 AGTAAAAATTAAATTTTCTAGATCCACTTCTAGGAAAACTATTAGAGTACTGGTTTTTA

a: SerPheLeuIleEndLysAspLeuGlyGluAspProPheEndEndSerHisAspGlnAsn -
 b: HisPheEndPheLysArgIleEndValLysIleLeuPheAspAsnLeuMetThrLysIle -
 c: IlePheAsnLeuLysGlySerArgEndArgSerPheLeuIleIleSerEndProLysSer -

3121 CCCTTAACGTGAGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATC
-----+-----+-----+-----+-----+-----+ 3180
GGGAATTGCACTCAAAGCAAGGTGACTCGCAGTCTGGGGCATCTTTTCTAGTTTCCTAG
a: ProLeuThrEndValPheValProLeuSerValArgProArgArgLysAspGlnArgIle -
b: ProEndArgGluPheSerPheHisEndAlaSerAspProValGluLysIleLysGlySer -
c: LeuAsnValSerPheArgSerThrGluArgGlnThrProEndLysArgSerLysAspLeu -
3181 TTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACACCGCT
-----+-----+-----+-----+-----+-----+ 3240
AAGAACTCTAGGAAAAAAGACGCGCATTAGACGACGAACGTTTGTTTTTTTGGTGGCGA
a: PheLeuArgSerPhePheSerAlaArgAsnLeuLeuLeuAlaAsnLysLysThrThrAla -
b: SerEndAspProPhePheLeuArgValIleCysCysLeuGlnThrLysLysProProLeu -
c: LeuGluIleLeuPhePheCysAlaEndSerAlaAlaCysLysGlnLysAsnHisArgTyr -
3241 ACCAGCGGTGGTTTTGTTTGGCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGG
-----+-----+-----+-----+-----+-----+ 3300
TGGTCGCCACCAAACAAACGGCCTAGTTCTCGATGGTTGAGAAAAAGGCTTCCATTGACC
a: ThrSerGlyGlyLeuPheAlaGlySerArgAlaThrAsnSerPheSerGluGlyAsnTrp -
b: ProAlaValValCysLeuProAspGlnGluLeuProThrLeuPheProLysValThrGly -
c: GlnArgTrpPheValCysArgIleLysSerTyrGlnLeuPhePheArgArgEndLeuAla -
3301 CTTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCA
-----+-----+-----+-----+-----+-----+ 3360
GAAGTCGTCTCGCGTCTATGGTTTATGACAGGAAGATCACATCGGCATCAATCCGGTGGT
a: LeuGlnGlnSerAlaAspThrLysTyrCysProSerSerValAlaValValArgProPro -
b: PheSerArgAlaGlnIleProAsnThrValLeuLeuValEndProEndLeuGlyHisHis -
c: SerAlaGluArgArgTyrGlnIleLeuSerPheEndCysSerArgSerEndAlaThrThr -
A
I
W
N
I
3361 CTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGC
-----+-----+-----+-----+-----+-----+ 3420
GAAGTTCTTGAGACATCGTGGCGGATGTATGGAGCGAGACGATTAGGACAATGGTCACCG
a: LeuGlnGluLeuCysSerThrAlaTyrIleProArgSerAlaAsnProValThrSerGly -
b: PheLysAsnSerValAlaProProThrTyrLeuAlaLeuLeuIleLeuLeuProValAla -
c: SerArgThrLeuEndHisArgLeuHisThrSerLeuCysEndSerCysTyrGlnTrpLeu -
3421 TGCTGCCAGTGGCGATAAGTCGTGCTTACCGGGTTGGACTCAAGACGATAGTTACCGGA
-----+-----+-----+-----+-----+-----+ 3480
ACGACGGTCACCGCTATTTCAGCACAGAATGGCCCAACCTGAGTTCTGCTATCAATGGCCT
a: CysCysGlnTrpArgEndValValSerTyrArgValGlyLeuLysThrIleValThrGly -
b: AlaAlaSerGlyAspLysSerCysLeuThrGlyLeuAspSerArgArgEndLeuProAsp -
c: LeuProValAlaIleSerArgValLeuProGlyTrpThrGlnAspAspSerTyrArgIle -

TAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAAC
3481 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3540
ATTCCGCGTCGCCAGCCCGACTTGCCCCCAAGCACGTGTGTGGGGTCGAACCTCGCTTG

a: EndGlyAlaAlaValGlyLeuAsnGlyGlyPheValHisThrAlaGlnLeuGlyAlaAsn -
b: LysAlaGlnArgSerGlyEndThrGlyGlySerCysThrGlnProSerLeuGluArgThr -
c: ArgArgSerGlyArgAlaGluArgGlyValArgAlaHisSerProAlaTrpSerGluArg -

GACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGA
3541 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3600
CTGGATGTGGCTTGACTCTATGGATGTGCGACTCGATACTCTTTCGCGGTGCGAAGGGCT

a: AspLeuHisArgThrGluIleProThrAlaEndAlaMetArgLysArgHisAlaSerArg -
b: ThrTyrThrGluLeuArgTyrLeuGlnArgGluLeuEndGluSerAlaThrLeuProGlu -
c: ProThrProAsnEndAspThrTyrSerValSerTyrGluLysAlaProArgPheProLys -

AGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCAGGAG
3601 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3660
TCCCTCTTTCGCGCTGTCCATAGGCCATTCGCGCTCCAGCCTTGTCTCTCGCGTGCTC

a: ArgGluLysGlyGlyGlnValSerGlyLysArgGlnGlyArgAsnArgArgAlaHisGlu -
b: GlyArgLysAlaAspArgTyrProValSerGlyArgValGlyThrGlyGluArgThrArg -
c: GlyGluArgArgThrGlyIleArgEndAlaAlaGlySerGluGlnGluSerAlaArgGly -

GGAGCTTCCAGGGGAAACGCCTGGTATCTTTATAGTCCTGTGCGGGTTTCGCCACCTCTG
3661 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3720
CCTCGAAGGTCCCCCTTTGCGGACCATAGAAATATCAGGACAGCCCAAAGCGGTGGAGAC

a: GlyAlaSerArgGlyLysArgLeuValSerLeuEndSerCysArgValSerProProLeu -
b: GluLeuProGlyGlyAsnAlaTrpTyrLeuTyrSerProValGlyPheArgHisLeuEnd -
c: SerPheGlnGlyGluThrProGlyIlePheIleValLeuSerGlyPheAlaThrSerAsp -

ACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAG
3721 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3780
TGAACCTCGCAGCTAAAAACACTACGAGCAGTCCCCCGCCTCGGATACCTTTTTGCGGTC

a: ThrEndAlaSerIlePheValMetLeuValArgGlyAlaGluProMetGluLysArgGln -
b: LeuGluArgArgPheLeuEndCysSerSerGlyGlyArgSerLeuTrpLysAsnAlaSer -
c: LeuSerValAspPheCysAspAlaArgGlnGlyGlyGlyAlaTyrGlyLysThrProAla -

CAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATG
3781 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3832
GTTGCGCCGGAATAAGCAAGGACCGGAAAACGACCGGAAAACGAGTGTAC

a: GlnArgGlyLeuPheThrValProGlyLeuLeuLeuAlaPheCysSerHis -
b: AsnAlaAlaPheLeuArgPheLeuAlaPheCysTrpProPheAlaHisMet -
c: ThrArgProPheTyrGlySerTrpProPheAlaGlyLeuLeuLeuThr -

Figur 9

RBS XbaI RBS
 CACAGGAAACAGCTATGACCATGATTACGAATTTCTAGATAACGAGGGCAAAAAATGAAAAAGACAGCTATCGCG
 LacZ: MetThrMetIleThrAsnPheEnd OmpA: MetLysLysThrAlaIleAla

StuI BsaI EcoRI SstI KpnI SmaI BamHI
 ATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAGGCCGAGACCAGAATTCGAGCTCGGTACCCGGGGAT
 IleAlaValAlaLeuAlaGlyPheAlaThrValAlaGlnAlaEnd

XhoI SalI PstI Eco47III HindIII
 CCCTCGAGGTCGACCTGCAGGCagcGCTTGGCGTCACCCGAGTTCGGTGGTTAATAAGCTTGACCTGTGAAGTG
 Strep-tag: SerAlaTrpArgHisProGlnPheGlyGlyEnd